

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:12:37 ; Search time 46 Seconds  
(without alignments)  
3861.668 Million cell updates/sec

Title: US-09-981-947B-2  
Perfect score: 2903  
Sequence: 1 MARKQNRNSKELGLVPLTDD.....NTAWIQAAATTTATKKEHPL 563

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 31518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_invertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2463	84.8	511	4 Q8WV30	Q8WV30 homo sapien
2	2242	77.2	538	11 Q7TSP0	Q7TSP0 mus musculus
3	1713.5	59.0	431	11 Q7TSK9	Q7TSK9 mus musculus
4	1681	57.9	416	11 Q9J3J2	Q9J3J2 mus musculus
5	1119.5	38.6	501	6 Q863Y9	Q863Y9 equus caball
6	1112.5	38.3	501	11 Q8RIN7	Q8RIN7 mus musculus
7	1101.5	37.9	501	6 Q8WMN1	Q8WMN1 ovis aries
8	1006	34.7	199	4 Q96P00	Q96P00 homo sapien
9	1001.5	34.5	499	4 Q8WYM3	Q8WYM3 homo sapien
10	999	34.4	496	4 Q8WYM5	Q8WYM5 homo sapien
11	999	34.4	503	4 Q8WYM4	Q8WYM4 homo sapien
12	994	34.2	503	4 Q8WYF9	Q8WYF9 homo sapien
13	914.5	31.5	220	11 Q8BZK3	Q8BZK3 mus musculus
14	899	31.0	489	13 Q8WSW5	Q8WSW5 eptatretus
15	895	30.8	492	13 Q918M1	Q918M1 oncorhynchus
16	894	30.8	505	13 Q8JFQ9	Q8JFQ9 oncorhynchus

17	884.5	30.5	493	13 Q801M9	Q801M9 xenopus lae
18	878.5	30.3	553	5 Q8MXW2	Q8MXW2 halocynthia
19	878	30.2	503	13 Q9DFV2	Q9DFV2 salmo trutt
20	875	30.1	509	6 Q8WJ15	Q8WJ15 equus caball
21	851	29.3	481	5 Q9V409	Q9V409 drosophila
22	847	29.2	478	13 Q918L9	Q918L9 cyprinus ca
23	839.5	28.9	493	11 Q8BLF7	Q8BLF7 mus musculu
24	829.5	28.6	497	4 Q8TDB9	Q8TDB9 homo sapien
25	828	28.5	520	4 Q8TDB8	Q8TDB8 homo sapien
26	785	27.0	533	13 Q7ZT03	Q7ZT03 ctenopharyn
27	770	26.5	396	11 Q8BTM5	Q8BTM5 mus musculu
28	710	24.5	482	13 Q98TW6	Q98TW6 oncorhynchus
29	703	24.2	510	5 Q44827	Q44827 caenorhabdi
30	697	24.0	492	5 Q44826	Q44826 caenorhabdi
31	687	24.0	500	5 Q27079	Q27079 taenia soli
32	664	22.9	516	5 Q44616	Q44616 caenorhabdi
33	663.5	22.9	492	5 Q26580	Q26580 schistosoma
34	648	22.3	507	5 Q9V310	Q9V310 drosophila
35	646	22.3	521	5 Q26579	Q26579 schistosoma
36	600	20.7	600	5 Q95TE8	Q95TE8 drosophila
37	600	20.7	647	5 Q8IRI6	Q8IRI6 drosophila
38	582	20.0	505	5 Q26581	Q26581 schistosoma
39	579	19.9	510	5 Q27072	Q27072 taenia soli
40	557.5	19.2	491	5 Q9U622	Q9U622 drosophila
41	547.5	18.9	476	5 Q8T3V4	Q8T3V4 drosophila
42	542.5	18.7	476	5 Q9V309	Q9V309 drosophila
43	530.5	18.3	516	5 Q9V919	Q9V919 drosophila
44	529	18.2	586	5 Q8MLN1	Q8MLN1 drosophila
45	524	18.1	546	10 Q93Z41	Q93Z41 arabidopsis

#### RESULT 1

Q8WV30	PRELIMINARY;	PRT;	511 AA.
ID	Q8WV30		
AC	Q8WV30;		
DT	01-MAR-2002 (TREMELrel. 20, Created)		
DT	01-MAR-2002 (TREMELrel. 20, Last sequence update)		
DT	01-OCT-2003 (TREMELrel. 25, Last annotation update)		
DE	Similar to solute carrier family 2 (Facilitated glucose transporter), member 9.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Ovary;		
RA	Strausberg R.;		
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.		
CC	-1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.		
DR	EMBL; BC018897; AAH18897.1; -.		
DR	GO; GO:0016021; C:integral to membrane; IEA.		
DR	GO; GO:0005351; F:sugar porter activity; IEA.		
DR	GO; GO:0005215; F:transporter activity; IEA.		
DR	GO; GO:0008643; P:carbohydrate transport; IEA.		
DR	InterPro; IPR07114; MFS.		
DR	InterPro; IPR005828; Sub transporter.		
DR	InterPro; IPR003663; Sugar transport.		
DR	InterPro; IPR005829; Sug transporter.		
DR	Pfam; PF00083; sugar tr; 1.		
DR	PRINTS; PR00171; SUGTRNSPORT.		
DR	TIGRFAMS; TIGR00879; SP; 1.		
DR	PROSITE; PS00850; MFS; 1.		
DR	PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.		
DR	PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.		
KW	Sugar transport; Transmembrane; Transport.		
SQ	SEQUENCE 511 AA; 55737 MW; C4BCB7138BA0F01E CRC64;		

Query Match 84.8%; Score 2463; DB 4; Length 511;



248	Qy	SLPFLPDS	PRVLLLEKHNEAR	KA	FO	FLG	KAD	VSQ	VEE	VLAE	SHV	QR	S	IR	L	V	S	V	L	E		300
209	Db	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	208	
308	Qy	LRAPV	RQVV	TV	IT	MAC	Y	QL	CL	GL	NA	I	FW	Y	NS	I	FG	K	AG	I	PP	367
209	Db	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	241	
368	Qy	AVFSG	LV	EH	GRR	PL	LG	FG	LM	GL	FG	T	L	T	I	T	L	Q	H	A	P	427
242	Db	AI	FSG	LV	ER	GRR	PL	GG	FM	AL	FG	T	L	T	L	T	L	Q	A	P	V	301
428	Qy	SGPG	GI	PI	L	GE	FP	Q	Q	S	O	R	P	A	A	I	I	A	G	V	N	487
302	Db	SGPG	GI	PI	L	GE	FP	Q	Q	S	O	R	P	A	A	I	I	A	G	V	N	361
488	Qy	CITGA	Y	I	V	F	V	L	P	E	T	K	N	T	Y	A	B	I	S	A	P	546
362	Db	CIAG	T	Y	F	V	L	P	E	T	K	N	T	Y	A	B	I	S	A	P	K	419

```

RESULT 4
Q99JJ2          PRELIMINARY;      PRT;   416 AA.
AC  Q99JJ2;
DT  01-JUN-2001 (TrEMBLrel. 17, Created)
DT  01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT  01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE  Hypothetical protein.
DS
OS  Mus musculus (Mouse).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
CX  NCBI_TaxID=10090;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  Strausberg R.;
RL  Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC  -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC  -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR  EMBL; BC008076; AAH0076.1; -.
DR  MGD; MGI:2152844; SLC2A9.
DR  GO; GO:0016021; C:integral to membrane; IEA.
DR  GO; GO:0003351; F:sugar porter activity; IEA.
DR  GO; GO:0005215; F:transporter activity; IEA.
DR  GO; GO:0006810; P:transport; IEA.
DR  InterPro; IPR007114; MFS.
DR  InterPro; IPR005828; Sub_transporter.
DR  InterPro; IPR005829; Sug_transporter.
DR  Pfam; PF00083; sugar tr; 1.
DR  PROSITE; PS00850; MFS; 1.
DR  PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR  PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR  Hypothetical protein; Sugar transport; Transmembrane; Transport.
SQ  SEQUENCE 416 AA; 45096 MW; 0A72AAAE5BF7AF2B CRC64;

```

	Query Match	57.9%;	Score 1681;	DB 11;	Length 416;	
	Best Local Similarity	67.7%;	Pred. No. 6.5e-115;			
	Matches 337;	Conservative	26;	Mismatches 27;	Indels 108;	Gaps 2
QY	49	RKDWCSLLVASLAGAGSGSFLGYNLSVVVNATPTPKAFYNESWERRHGRIPDPTLTL	108			
		:::::				
Db	15	QRKWFSLVVAALVGAFGSSFLGYNLSVVVNATPTPKAFYNGTWYRRHQPIDPDTLTL	74			
QY	109	LMSVTYSIIFAIGSLGVTLLIVNMIGKVLRGRKHITLLANNGFPAISALLMACSLQAGPEMLI	168			
		:::::				
Db	75	LMSVTYSIIFAIGSLGVTLMVMYMGFKPLGRKSTILLVNNGFPAISALLMACSLRAGTPEMLI	134			
QY	169	VGRFIMGIDGGVALSVLPWYLSEISPKERGSGLQVTAIFICIGVTGQLLGPPELLGKE	228			
		:::::				
Db	135	VGRFIMGVDGGIALSALPMWLNEISPKEIRGSGLQVTAIFICIGVSQQLLGPPELLGR-	193			
QY	229	STWPYLFGVIVPAVVQLLSLPELDPSRYVLLEKNEARAVKAQTFFISKADVSQVEE	268			

QY 56 LLVASLAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRIPDDPTLLWSVTVS 115  
DB 17 LALATLIAAFSSFOGYNVAVNSPAELKAFYNETHYSRFSYSEFSLTLLWSVVS 76  
QY 116 IFATGGLVGTLLVGMKVLGRKHTLLANNGFSAALLMACSLQAGAFEMLIYGRFTMG 175  
DB 77 MFPFGFVGLMWGLVNLGRKGTLLFNNIFSVIPALMTGSKTARSYEMIIISRLLVG 136  
QY 176 IDGVALSVLPMYLSIEISPKIRSGVGTAFICIGVFTGQLLGLPELLGKSTWPLYF 235  
DB 137 ICAGLSSNVPMYLGELSPKLRGALGVVPQLFITVGLVLAQIVGLRSLLTATESGPILL 196  
QY 236 GVIVPAVQVLLSPFLPDSRYLLEKKNBARAVKAFQTLGKADVSQEVVEVLASHV 295  
DB 197 GLTAIPALQQLLLFPFSPRYLLIQKDAAKKALKRLRGWDDVDVAEMEEIQLEDEA 256  
QY 296 QRSIRLVSVLELLRAPYVRQVVTVMACVOLGCLNAINFYNSIFGKAGIPPAKIPY 355  
DB 257 EKAAGTISVLTFRVSRSLRWQVSIILMGGQQLSGVNAIYYADQIYLSAGVKDQDVQY 316  
QY 356 VTLSGGTETLAAVFSGLVIEHLGRPELLIGSGLMGLFFGTLLTTLTLODHAPVWPYLS 415  
DB 317 VTVGTGANVMTICAVFVVEYLGRRALLLGLFSCVFIACCVLTVALAQDRVSMPYIS 376  
QY 416 IVGILAIASFCSGPGGIPFIITGTFQOSQRPAAFIAGTVNLSNFAVGLLFPFIQKS 475  
DB 377 IVCVSVYVIGHALGSPSPALLITEVFLQSSRAAYMVGTVHMLSNFAVGLVFPFIQVG 436  
QY 476 LDTYCFVLPATICITGAIYLYFVLBPETKNRYVAEISQAFSKNK---AYPPEEKIDSMT 532  
DB 437 LGAYSFIIPAVICLLTIYFVLVBPETKGTVEINHIPTKMKVSDVHPAKDEL----K 492  
QY 533 DAPASS 538  
DB 493 DIPLSA 498

RESULT 6

Q8RWNI PRELIMINARY; PRT; 501 AA.  
ID Q8RWNI7;  
AC Q8RWNI7;  
DT 01-JUN-2002 (TrEMBLrel. 21, Created)  
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Similar to solute carrier family 2 (Facilitated glucose transporter), member 5.  
GN SLC2A5.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA Strausberg R.;  
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Testis;  
RX MEDLINE=22354683; PubMed=12466851;  
RA The PANTOM Consortium,  
RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs."  
RT Nature 420:563-573(2002).  
RL Nature 420:563-573(2002).  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
DR EMBL; BC023500; AAH23500.1; -.  
DR EMBL; AK029720; BAC26582.1; -.  
DR MGD; MGI:1928369; SLC2A5.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005351; F:sugar porter activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.

DR GO; GO:000643; P:carbohydrate transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR InterPro; IPR003663; Sugar\_transpt.  
DR InterPro; IPR005829; Sug\_transporter.  
DR Pfam; PF00083; sugar tr; 1.  
DR PRINTS; PR00171; SUGRTNSPORT.  
DR TIGRfams; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Sugar transport; Transmembrane; Transport.  
SQ SEQUENCE 501 AA; 55409 MW; 4F562A3BF848B5C7 CRC64;  
Query Match 38.3%; Score 1112.5; DB 11; Length 501;  
Best Local Similarity 45.8%; Pred. No. 3.4e-73;  
Matches 217; Conservative 94; Mismatches 160; Indels 3; Gaps 1;  
QY 56 LLVASLAGAFSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRIPDDPTLLWSVTVS 115  
DB 16 LALATLIAAFSSFOGYNVAVNSPAELKAFYNETHYSRFSYSEFSLTLLWSVVS 75  
QY 116 IFATGGLVGTLLVGMKVLGRKHTLLANNGFSAALLMACSLQAGAFEMLIYGRFTMG 175  
DB 76 MFPFGFVGLMWGLVNLGRKGTLLFNNIFSVIPALMTGSKTARSYEMIIISRLLVG 135  
QY 176 IDGVALSVLPMYLSIEISPKIRSGVGTAFICIGVFTGQLLGLPELLGKSTWPLYF 235  
DB 136 ICAGLSSNVPMYLGELSPKLRGALGVVPQLFITVGLVLAQIVGLRSLLTATESGPILL 195  
QY 236 GVIVPAVQVLLSPFLPDSRYLLEKKNBARAVKAFQTLGKADVSQEVVEVLASHV 295  
DB 196 GLTGPAGLQQLLLFPFSPRYLLIQKDEAAERALQTLRGKDVHLEMEIRKEDEA 255  
QY 296 QRSIRLVSVLELLRAPYVRQVVTVMACVOLGCLNAINFYNSIFGKAGIPPAKIPY 355  
DB 256 EKAAGTISVLTFRVSRSLRWQVSIILMGGQQLSGVNAIYYADQIYLSAGVKDQDVQY 315  
QY 356 VTLSGGTETLAAVFSGLVIEHLGRPELLIGSGLMGLFFGTLLTTLTLODHAPVWPYLS 415  
DB 316 VTVGTGANVMTICAVFVVEYLGRRALLLGLFSCVFIACCVLTVALAQDRVSMPYIS 375  
QY 416 IVGILAIASFCSGPGGIPFIITGTFQOSQRPAAFIAGTVNLSNFAVGLLFPFIQKS 475  
DB 376 IVCVSVYVIGHALGSPSPALLITEVFLQSSRAAYMVGTVHMLSNFAVGLVFPFIQVG 435  
QY 476 LDTYCFVLPATICITGAIYLYFVLBPETKNRYVAEISQAFSKNK---AYPPEEK 526  
DB 436 LGAYSFIIPAVICLLTIYFVLVBPETKGTVEINHIPTKMKVSDVHPAKDEL----K 489  
RESULT 7  
Q8RWNI PRELIMINARY; PRT; 501 AA.  
ID Q8RWNI7;  
AC Q8RWNI7;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glucose transporter 5.  
GN SLC2A5.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Small intestine;  
RA Wood I.S.; Castano-Wereditz E.F.; Dyer J.; Shirazi-Beechey S.P.;  
RT "Nutrient regulation of Gluts in sheep intestine."  
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.



DR EMBL, AJ315928; CAC86964.1; ...  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005622; C: intracellular; IEA.  
DR GO; GO:0005351; F: sugar porter activity; IEA.  
DR GO; GO:0003700; F: transcription factor activity; IEA.  
DR GO; GO:0005215; F: transporter activity; IEA.  
DR GO; GO:0008643; P: carbohydrate transport; IEA.  
DR GO; GO:0008355; P: regulation of transcription, DNA-dependent; IEA.  
DR InterPro; IPR000005; HTHRAC.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub transporter.  
DR InterPro; IPR003663; Sugar transp.  
DR InterPro; IPR005829; Sug transporter.  
DR Pfam; PF00083; sugar tr; 1.  
DR PRINTS; PR00171; SUGTRNSPORT.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00041; HTH ARAC\_FAMILY\_1; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
DR SUGAR transport; Transmembrane; Transport.  
KW SUGAR transport; Transmembrane; Transport.  
SQ SEQUENCE 501 AA; 55546 MW; 23F0B43E944C9AD5 CRC64;  
  
Query Match 37.9%; Score 1101.5; DB 6; Length 501;  
Best Local Similarity 45.7%; Pred. No. 2.2e-72;  
Matches 217; Conservative 97; Mismatches 158; Indels 3; Gaps 1;  
  
QY 56 LLVASLAGAGSFLXYGNLSVNVNAPTYKAFYNESWERRGRPDPTLLLSVTVS 115  
DB 17 IVLATLAFGSGFQGVNVATINSSEFWKPFYNTYDVRGEYNEVYLLLSVTVS 76  
  
QY 116 IFAIGLVGTLIYKMGKVLGRKHTLLANGFAISAAALLMACSLOQAFEMIVGRFIMG 175  
DB 77 MFPPGGFLGSLMVGPLVNLGRKGTLLFNINFSIVPALLMGFSDLAKSPFEMIIVARVLVG 136  
  
QY 176 IDGVALSVLPMYLSISPKIIRGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWPLYF 235  
DB 137 ICAGLSNVVPMYVLGSLAPKRWGALGVWPQFITTIGILVAQIFGRLSLANEEGPILL 196  
  
QY 236 GVIVPAVQVQLLSLPFDPSPRYLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHV 295  
DB 197 GLTGIPAVQLLFLPFPSPFRVLLIQKDEEAKRALRLRGWHVDVAIEIILEEDRA 256  
  
QY 296 QRSIRLVSVLELRAPYRQVIVTWACYQCLGNALWFYNTSIFGKAGIPPAKIPY 355  
DB 257 EXANGFISVLKFKMSRLRQVSIIVLMAGQLSGVNAIYVYADQIYLSAGVKEDDVQY 316  
  
QY 356 VTLSTGIEFLAAVFGSLVIEHGRPLIGGFLGMLPFGTLITLTLODHAPVPLYLS 415  
DB 317 VTAGTGA NVNLITVCAIFVVELMGRFRLLLLGFSVCFACCVLTGALAQDVISWMPYVS 376  
  
QY 416 IVCILAILIASFCSPGPGIPILLTGEFPQSQRPAAFIILAGTWNLSNFAVGLLPFIQKS 475  
DB 377 IACVISYVIGHALGPSPIPALLVTEIFLOSSRPAAYVAGTVHLSNFTVGLVFPPIQVG 436  
  
QY 476 LDYCYFLVAFATIGTGAIVLYFVLPETQNTYAEISQAFSRKNK--AYPPEKI 527  
DB 437 LGAYSFVFAVCLTIVTIVFLIIPETKSKTFIEINQIFIKMKVFGVHPEKEEL 491  
  
RESULT 8  
Q96P00 PRELIMINARY; PRT; 199 AA.  
AC Q96P00;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Solute carrier family 2 member 9 (Fragment).  
GN SUC2A9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
RC TISSUE=Articular cartilage;  
RA Neema G., Richardson S., Bell S., Carter S., Mobasheri A.;  
RT "Molecular characterization and cloning of glucose transporters in human articular chondrocytes";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
DR EMBL; AF421859; AAL16939.1; -  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005215; F: transporter activity; IEA.  
DR GO; GO:0006810; P: transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub transporter.  
DR Pfam; PF00083; sugar tr; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Transmembrane.  
FT NON\_TER 1  
FT NON\_TER 199  
SQ SEQUENCE 199 AA; 21541 MW; 17C0770C9B36F338 CRC64;  
  
Query Match 34.7%; Score 1006; DB 4; Length 199;  
Best Local Similarity 100.0%; Pred. No. 6.9e-66;  
Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 85 IKAFYNESWERRGRPDPTLLLSVTVSIFAIIGLVGTLIYKMGKVLGRKHTLLAN 144  
DB 1 IKAFYNESWERRGRPDPTLLLSVTVSIFAIIGLVGTLIYKMGKVLGRKHTLLAN 60  
  
QY 145 NGPAISAAALLMACSLOQAFEMIVGRFIMGIDGVALSVLPMYLSISPKIIRGSLGQV 204  
DB 61 NGPAISAAALLMACSLOQAFEMIVGRFIMGIDGVALSVLPMYLSISPKIIRGSLGQV 120  
  
QY 205 TAIPIICIGVFTGQLLGPPELLGKSTWPLYFGVIVPAVQVQLLSLPFDPSPRYLLEKH 264  
DB 131 TAIPIICIGVFTGQLLGPPELLGKSTWPLYFGVIVPAVQVQLLSLPFDPSPRYLLEKH 180  
  
QY 265 NEARAVKAFOTFLGKADVS 283  
DB 181 NEARAVKAFOTFLGKADVS 199  
  
RESULT 9  
Q8WYM3 PRELIMINARY; PRT; 499 AA.  
AC Q8WYM3;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Facilitative glucose transporter GLUT11-c.  
GN SLC2A11-c.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]\_TaxID=9606;  
RP SEQUENCE FROM N.A.  
RA Shimizu N., Minoshima S., Sasaki T.;  
RT "Molecular cloning of a novel member of the glucose transporter family";  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
DR EMBL; AB067444; BAB83506.1; -  
DR GO; GO:0016021; C: integral to membrane; IEA.  
DR GO; GO:0005351; F: sugar porter activity; IEA.  
DR GO; GO:0005215; F: transporter activity; IEA.  
DR GO; GO:0008643; P: carbohydrate transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub transporter.  
DR InterPro; IPR003663; Sugar transp.

```

DR      InterPro; IPR005928; Sub_transporter.
DR      InterPro; IPR003683; Sugar_transporter.
DR      InterPro; IPR005829; Sugar_transporter.
DR      Pfam; PF00083; sugar tr; 1.
DR      PRINTS; PR00171; SUGSTRNSPORT.
DR      TIGRFAMs; TIGR00879; SP; 1.
DR      PROSITE; PS50850; MFS; 1.
DR      PROSITE; PS00317; SUGAR_TRANSPORT_2; 1.
DR      Sugar transport; Transmembrane; Transport.
DR      SEQUENCE 496 AA; 53674 MW; 5500D9E2CC6A2ED6D CRC64;
SQ

```

Query Match 34.4%; Score 999; DB 4; Length 496;  
Best Local Similarity 42.1%; Pred. No. 6.8e-65;  
94; Mismatches 174; Indels 0; Gaps 0;

56	LLVASLAGAGSSFLYGVNLSVNAPTPIKAFYNESWERRRGRPDPTLLTLNWSVTYS	115	
Db	13	LLTUCAGIGGTFQYGNLSIINAPTLLHQFTNETWQARTGEPLDHLVLLMWSLIYS	72
Qy	116	IPAIGLGLVGLTIKVMKIGKVLGRKHTLLANNGFAISAALLMACSLQAGFEMLIVGRFIMG	175
Db	73	LYPLGGLFGALLAGPLAITLGRKSKLLVNNIFVWSAAILFGFSRKAGSFEMIMLGRLLVG	132
Qy	176	IDGVALSVLPWYLSISPKETRSLGQVTAIFCIGVFTGQLLGLPELLGKESTWPYLYF	235
Db	133	VNAGVSNMTIQPYWLGESAPKELGAVMSAIFTALGIVNGQVVGRELLGGQANPLLL	192
Qy	236	GVTVPVAVVOLLSPPLPDSPRYLLLEKHEARAVKAFOTFLGKADVSQVEVSVLAESHV	295
Db	193	ASCLVPGALQALASLLPLPSPRYLLDCGDTEACLAALRQLRGSGDLAGELEBERAA	252
Qy	296	QRSTRLVSVLELLRAPYVRQVVTVIWMACYQCLGNIAFWTNSIFGKAGTTPPAKIPY	355
Db	253	CQCRARRPWELFQHRALRRQVTSVLVLSGAMBELCGNDSVYAYASSVFRKAGVPEAKIQY	312
Qy	356	VTLSGTGIEFLAAVFSGLVIEHGRPLLLGGFLAGLFFGTITITLTLQDHAPWVPYLS	415
Db	313	AIIGTGSCELLTAWSCWJIERVGRVLLIGGYSLMTGWSIFTFVALCLASSPFWTYLA	372
Qy	416	IVGLIAIASFSGPGGIPETLAGEFPOQQRPAAFIAGTVNWLNSNPAVLGLPFPFOKS	475
Db	373	MACIFAFILSFGIGPAGVTGILATELFDQMARPAACWCGALNMIMLVLGLGFPFIMEA	432
Qy	476	LDTYCFLVPATICITGAIYIVFYLPETKNTYAEISQAFKRN	518
Db	433	LSHFYVPPFLGVCCGAIYVGLFPEPKGTGTFOBISXELHRLN	475

RESULT 11

```

Q8WMY4
ID Q8WMY4 PRELIMINARY; PRT; 503 AA.
AC Q8WMY4
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Facilitative glucose transporter GLUT11-b.
GN SLC2A11-B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shimizu N., Minoshima S., Sasaki T.;
RT "Molecular cloning of a novel member of the glucose transporter
RT family."
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AB067443; BAB83505.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.

```

[illegible]

```

RESULT 12
Q8WXKF9
ID Q8WXF9 PRELIMINARY; PRT; 503 AA.
AC Q8WXF9;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Sugar transporter GLUT11 long form.
DE OS Homo sapiens (Human).
DE Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
DE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
DE NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RT Wu X., Li W., Freeze H.H.;
RT "Cloning and Characterization of GLUT11, a Novel Sugar Transporter
RT with Alternative Splicing Forms.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DDBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.
DR EMBL; AF443201; AAL39061.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.

```

DR	GO: 0003215; F: transporter activity; IEA.	
DR	GO: 0008643; P: carbohydrate transport; IEA.	
DR	InterPro: IPR007114; MFS.	
DR	InterPro: IPR005828; Sub: transporter.	
DR	InterPro: IPR003663; Sugar transport.	
DR	InterPro: IPR005829; Sug: transporter.	
DR	Pfam: PF00683; sugar trn_1.	
DR	PRINTS; PR00171; SUGTRNSPORT.	
DR	TIGRFAMS; TIGR00879; SP; 1.	
DR	PROSITE; PS0850; MFS; 1.	
DR	PROSITE; PS02117; SUGAR TRANSPORT 2; 1.	
KW	Sugar transport; Transmembrane; Transport.	
SQL	SEQUENCE 503 AA; 54545 MW; 2C682EAA3D7BB889 CRC64;	

  

Query Match	34.2%;	Score 994;	DB 4;	Length 503;
Best Local Similarity	41.9%;	Pred. No. 1.6e-64;		
Matches 194;	Conservative 94;	Mismatches 175;	Indels 0;	Gaps 0;

  

Qy	56	LLVASLAGPSSFLYCYNLSVVNAETPIKAFYNESWERRHGRDPDPTLLTLKMSVTS	115
Db	20	LLLTICAAIGIGTFOFGYCNLSINAPTLHIQEFNTWQARTGEPDHLVLLMSLVS	79
Qy	116	IFAIGGLVGLIVKMGKVLGRKHTLLANNNGFAISAAWMSLOQAGAFEMLVIGRFMG	175
Db	80	LYPLGLGFGALLAGPAATLGRKSLVNINIFVWSAAILFGPRKAGSFEMLMGRLVIG	139
Qy	176	IDGVALSVLPMLYSISPKIRGSLGQVTAIPICIGVFTGQLLGLPELLGKESTWPLYF	235
Db	140	VNAGVKNITOPMYLGSAPKELRGAVAMSAIFTALGIVMGVVGLRELLGGPQAWPELL	199
Qy	236	GVTWPAVQLLSLPFLPSRVLLEKNEARVAKPQTFLGKADVSDEVLEVAESHV	295
Db	200	ASCLVPGALQLASLPLLPSPRYLLIDCGTEACTAELRRLRGSGDLAGELEBEERAA	259
Qy	296	QRGIRLVSVLELLRAPYVRKQWTVVTWACYQLCGNALWFVTSIFGKAGIPPAKIPY	355
Db	260	QCQCRARRPWFQFHALRRQVTSVLVLSAMELCNDNVYAVASSVFRKAGVPEAKIQY	319
Qy	356	VTLSGTGIEFLAAVFGSLVTEHLGRPLIGFGMLGFPFGTLITITLTDHAPMPVYLS	415
Db	320	AIIGTSCCELLTAVSCVVIERVRRVLLIGVSLMTGWSIFTVALCLQSFPFTLYLA	379
Qy	416	IVGILAITAIFSCSPGGIPFILTGEFFQOSRPAAFIIAGTVNWSLNFVGLLFFPIQKS	475
Db	380	MACIFAFILSFQIGPAGVTGILATELFDQWARPAACWVGALMWIMLILVGLGFPFIMEA	439
Qy	476	LDTYCFILVFATICITGAIYLVFVLPETKQRTYAEISQAFSKRN	518
Db	440	LSHFLYVPFGVCVCAIYVGLFPLTKGTFOISKELHRLN	482

```

RESULT 13
Q8BZR3 Q8BZR3 PRELIMINARY; PET; 220 AA.
AC Q8BZR3;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Solute carrier family 2 (Fragment).
GN SUC2A9.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
SEQUENCE FROM N.A.
RP STRAIN=C57BL/6J; TISSUE=Cecum;
RC MEDLINE=22354683; PubMed=12466851;
RX The FANTOM Consortium.
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).

```

KW Sugar transport; Transmembrane; Transport.  
SQ SEQUENCE 489 AA; 53664 MW; 351F0B340E19B697 CRC64;  
Query Match 31.0%; Score 899; DB 13; Length 489;  
Best Local Similarity 39.7%; Pred. No. 1.4e-57;  
Matches 195; Conservative 84; Mismatches 200; Indels 12; Gaps 6;  
QY 48 RKDWSCSLVASLAGAGSSFLYGYNSVNVNAPPYIKAFYNESWERRHRPDIDILIT 107  
Db 4 KEKQLTFRLAFVTVVAAALG-SLLFGFNIGVWNAPEQIIKDFNFETWGRSGVIEIKATIL 62  
QY 108 LLWSVTVSIFAIGLVGLVILVIMKIGLGRKHTLLANNFAISALLACSLQAGAFEML 167  
Db 63 TLWSETVAIFTVGGVSLVGLFVNRFRNPMCLNLLAVLGGAFVAMSKDASSYETL 122  
QY 168 IVGRFMIGDGGVALSVLPYLSEISPKIRGSLQVTAIFICIGVFTGQLLGLPELLGK 227  
Db 123 ILGRFVIGYCGLATGTFVPMYLGELISPTNLRGAVGTIHLAIVFGILLISQVLGNFI 182  
QY 228 ESTWPLVFGVIVPAVQVLLSLPFLSDPSRYLLLEKHNARAVKAFQTFGLKADVSQVE 287  
Db 183 ENRWPIILLGLAIIPAVVQALALPFCPKSPRFLINQTEKEAKDLVKQLRGVEDVGTML 242  
QY 288 EVLAESHVORSIRLVSVLELLRAPVVRQVTVITMACYQLCGLNATWFTYNSIFGKAG 347  
Db 243 ENKEEHRMTQBPKVSIQLQFRHPNVRQAIISIVLSQLSGINAVFYVSTGIFSKAG 302  
QY 348 IPPAKIPYTLTGSTGIEHLAIVFSGVLVIEHLGRRLILGGGLMGLFGTTLITLQDH 407  
Db 303 VDQAE--YATIGAGVYNAFTIVSLFVERLGRRLHLVGLGGMAVCTVIMTSLUHEWTK 360  
QY 408 APWVPLSVIGILAIITASFCSGPGGIPFILTGEFFQSQORPAAFIAGTVNWSLNFAYGL 467  
Db 361 APAVSYLAIATFGFVAFVFEIGGIPWFIIVAEISQGPRAVAVAGCSNWTSLFLVAM 420  
QY 468 LFPFIQKSLDTYCLVFAFATIGTAYLYFVLVLPETKNRTYAEISQAFSKRKA-YPPE-- 524  
Db 421 LFPFAQELMGPPFVFLIFTVLLIFFTFTYFRVETPKRTFTDDIASEF--RVKAGMKPEAA 478  
QY 525 EKI-----DSAV 531  
Db 479 ENIQNTDSSV 489  
RESULT 15  
Q918M1 PRELIMINARY; PRT; 492 AA.  
AC Q918M1;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glucose transporter 1A.  
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
CX NCBI\_TaxID=8022;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20571038; PubMed=11121589;  
RA Teerijoki H., Krasnov A., Pitkanen T.I., Molsa H.;  
RT "Cloning and characterization of glucose transporter in teleost fish rainbow trout (Oncorhynchus mykiss).";  
RL Biochim. Biophys. Acta 1494:290-294(2000).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
DR EMBL; AF247728; AAF75681.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005351; F:sugar porter activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0008443; P:carbohydrate transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.

EMBL; AK033725; BAC28448.1; -.  
DR MGD; MGI:2152844; SLC2a9.  
DR GO; GO:0016020; C:membrane; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0006810; P:transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub transporter.  
DR InterPro; IPR005829; Sug\_transporter.  
DR Pfam; PF00083; sugar tr; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
FT NON TER 1  
SQ SEQUENCE 220 AA; 23861 MW; 033C140B030B86B9 CRC64;  
Query Match 31.5%; Score 914.5; DB 11; Length 220;  
Best Local Similarity 85.2%; Pred. No. 3.9e-59;  
Matches 178; Conservative 13; Mismatches 17; Indels 1; Gaps 1;  
QY 338 YTNISFGKAGIPPAKIPVYTLSTGGIETLAIVFSGVLVIEHLGRPLLIIGFGLMGLFFGT 397  
Db 1 YTNISFGKAGIPQKIPYTLSTGGIETLAIVFSGVLVIEHLGRPLLIIGFGLMGLFFGT 60  
QY 398 LTITLTLDQHPWPYLSIVGILAIITASFCSGPGGIPFILTGEFFQSQORPAAFIAGTV 457  
Db 61 LTATLTLDQHPWPYLSIVGILAIITASFCSGPGGIPFILTGEFFQSQORPAAFIAGTV 120  
QY 458 NWLNFVAGLLPFIQKSLDYCLVFAFATIGTAYLYFVLVLPETKNRTYAEISQAFSKR 517  
Db 121 NWLNFVAGLLPFIQKSLDYCLVFAFATIGTAYLYFVLVLPETKNRTYAEISQAFSKR 180  
QY 518 NKAYPPEKIDSAVTDAPASSPFTTPNTA 546  
Db 181 NKAQPEVKADSAVTEKANSQ-TEPDSS 208  
RESULT 14  
Q8WSW5 PRELIMINARY; PRT; 489 AA.  
AC Q8WSW5;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Glucose transporter.  
GN GLUT.  
OS Eptatretus stoutii (Pacific hagfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Hyperotreti; Myxiniiformes;  
OC Myxiniidae; Eptatretinae; Eptatretus.  
OX NCBI\_TaxID=7765;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RA Yao S.Y.M., Ng A.M.L., Baldwin S.A., Young J.D.;  
RT "Molecular cloning and functional characterization of an ancient pre-vertebrate glucose transporter (GLUT) from the Pacific hagfish (Eptatretus stoutii).";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY.  
DR EMBL; AY059413; AAL27090.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0005351; F:sugar porter activity; IEA.  
DR GO; GO:0005215; F:transporter activity; IEA.  
DR GO; GO:0008443; P:carbohydrate transport; IEA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR InterPro; IPR003663; Sugar transport.  
DR InterPro; IPR005829; Sug\_transporter.  
DR Pfam; PF00083; sugar tr; 1.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.

DR InterPro; IPR003663; Sugar\_transpt.  
DR InterPro; IPR005829; Sug\_transporter.  
DR Pfam; PF00083; sugar\_tr; 1.  
DR PRINTS; PR00171; SUGTRNSPORT.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Sugar transport; Transmembrane; Transport.  
SQ SEQUENCE 492 AA; 53687 MW; 39D13F9D6288AA56 CRC64;

Query Match 30.8%; Score 895; DB 13; Length 492;  
Best Local Similarity 37.8%; Pred. No. 2.8e-57;  
Matches 179; Conservative 100; Mismatches 187; Indels 8; Gaps 4;

QY	57	LVASLAGFGSSFLYGVNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLIWSVTVSI	116
Db	12	LMUTVGAAVIGSLQFGVNTGVINAPQKVIERFLNETWFDYQEPISKTSITLIWSVSVAI	71
QY	117	FAICGLVGTLLIVKMGVGRKHTLLANNGFAISAAALMACSLQAGAFEMLVGRFIMGI	176
Db	72	FSVGGIFGSPVGLFVNRFGRNRMELMANVLAFVSAALMGFSKVGSGWEMLIIGREVVGL	131
QY	177	DGVALSVLPMYLSISPKEIRSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLYFG	236
Db	132	YSGLSTGTFVPMYVGEVAPTALRGALGTLHQLGIVTGLMAQVFGMEALMGNASLWPFLLG	191
QY	237	VIIVPAVQVLLSLPELPDSPRYLLLEKHNEARAKAFQTLGKADVSQVEVEVLAESHVQ	296
Db	192	FTIPALAQCALLPFCPESPGFLINRNEENKAKTVLKLRTTVDYADMQEMKEARQM	251
QY	297	RSIRLYSVLELLRAPYVRWQVTVIVTMACYQLCGLNAINWFTYNSIFGKAGIPPAKIPYV	356
Db	252	MREKKVITILEFRSPLYRQPIFIAIMQLSQQLSGINAVFYIYSTRIFEXAGV-AQPVYA	309
QY	357	TLSTGTGTTAAVFSGLVIEHLGRPL-LIGGFLMGLPFGTLTITLTLODHAPWVPYLS	415
Db	310	TIGAGVVNTAFTVVSFLFVVERAGRSLLHLGLICMAGAAV-LMTIALALIDKLPWMSYVS	368
QY	416	IVGILAIISFCSPGGIPFILTGEFPQSQRPAAFIAGTVNVLNSNFAVGLLFPFIQKS	475
Db	369	IVAIFAFVAFPEIGPGPIPWFIVAELFSQGPSPSAFAVAGFSNWTANFIVGNAAFQYVEEL	428
QY	476	LDYCYFLVFATICTIGAILYLYFVLPETKNRYAEISQAFSKRN---KAYPPEE	525
Db	429	CGPYVVFVIFILLSPFIYFYFKVPETKGTDFDEISAGFRQSSGTGGEKHSPEE	482

Search completed: April 6, 2004, 10:18:26  
Job time : 48 secs

FT	DOMAIN	1	5	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	6	26	1 (POTENTIAL).
FT	DOMAIN	27	64	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	65	85	2 (POTENTIAL).
FT	DOMAIN	86	93	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	94	114	3 (POTENTIAL).
FT	DOMAIN	115	124	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	125	145	4 (POTENTIAL).
FT	DOMAIN	146	153	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	154	174	5 (POTENTIAL).
FT	DOMAIN	175	183	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	184	204	6 (POTENTIAL).
FT	DOMAIN	205	269	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	270	290	7 (POTENTIAL).
FT	DOMAIN	291	304	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	305	325	8 (POTENTIAL).
FT	DOMAIN	326	334	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	335	355	9 (POTENTIAL).
FT	DOMAIN	356	363	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	364	384	10 (POTENTIAL).
FT	DOMAIN	385	399	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	400	420	11 (POTENTIAL).
FT	DOMAIN	421	424	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	425	445	12 (POTENTIAL).
FT	DOMAIN	446	494	CYTOPLASMIC (POTENTIAL).
FT	SITE	277	279	DEFINES SUBSTRATE SPECIFICITY (BY
FT	CARBOHYD	43	43	SIMILARITY).
FT	SEQUENCE	494	AA; 54194	N-LINKED (GLCNAC... ) (POTENTIAL).
FT	SEQUENCE	494	AA; 54194	MM; A89204D3EA74BFBA CRC64;

Query Match 29.8%; Score 866; DB 1; Length 494;  
 Best Local Similarity 38.9%; Pred. No. 2.1e-51;  
 Matches 178; Conservative 90; Mismatches 188; Indels 2; Gaps 1;

QY	57	LVASLAGAGSGFLYGYNSVNNADPYIKAFYNESWERRGRDIDPDITLLMSVTVSI	116
DB	10	LFAISIAITIGSFQGYNTGVNAPEAIKDPFLNYLTERSETPPSSVLLTSLMSLSVAI	69
QY	117	FAIGGLVGLIIVKMGKVLGRKHTLLANNNGAISAAALMACSLOAGAFEMLIIVGRFIMGI	176
DB	70	FVGGMIGFSVGLFVNRFRGRNSMLIYNLAAGGCLMGFKIAESVEMLIIGRLIIGL	129
QY	177	DGQVALSVLPMYLSRISPKRISGSLGQVTAIFICIGVFTGQLGLPELLGKSTWPLYFG	236
DB	130	FCGLCTGFVPMYIGISPTALAGAGTLNQLGIVIGILVAQIFGLKVLGTEDLWPLLGG	189
QY	237	VIVPAVQLLSLPDPSRYLLEKNEARAVKAFOTFLGKADVQOEVEEVLAEASHVQ	296
DB	190	FTILPAIIQCAALPFCPSRFLINRKEEKAKEILQRLWGTEDVAQDIQEMKDESIRM	249
QY	297	RSIRLVSVLELRAPVVRQVVTIVTMACYOLCLNAIWFTNSISGKAGIPPAKIPYV	356
DB	250	SOEKQVTVLELFRAPNYRQPIIISIMQLSOLSGINAVFYISTGIFKADGV--QEPVYA	307
QY	357	TLSTGGIETLAARVFGSLVIEHLGRPLLIIGFGLMGLFFGTITITITLQDHAPWPVYLSI	416
DB	308	TIGAGVNTIFTVSVFLVERAGRTLHLIGLGNVAFCSILMTISLLKONYSNWSPICI	367
QY	417	VGILAIASFCGPGGIPILTGEFFQSQRPAAFIAGTWNLSNFAVGLLPPFIOKSL	476
DB	368	GAILVFAVFEIGPGPIPMFIVAELEFGQGPRAAMAVAGCSNWTSNFLVGLLFPSTFYL	427
QY	477	DTYCFVLPVATICITGAILYLYFVLPETKRTVAEISQAF	514
DB	428	GAYVIVFVFLVFWVFFFKVPETRGTEETRAP	465

Search completed: April 6, 2004, 10:17:27  
 Job time : 20 secs

352 KIPYVTLSTGGIETLAASFVSLVIEHLGRPL-LIGGFLGMLFFGTLTITLTLODHPW 410  
324 ---YATIGAGVNTVFTLVSLVLLVESAGRTTLHLGLAGMCGCAI-LMTVALLLDERVA 379  
411 VPYLSIVGILAIITASCSGPGGIPFLITGTEFFQSORPAAFIAGTVNLSNPVGLLFP 470  
380 MSYVSIVAIFFGFAFEIGPGIPFWFVAELFSQGPFPAMAVAGFSNNTSFIIGMGFQ 439  
471 FIOKSLDTYCLVLFATICTITGAIYLYFLVLPETKNRTYASISQAFSKRNKAYPPEEK 526  
440 YVRAEMGPYVLLFAVLLGLFFFTFLRVPETGRFTDQISAFHRTPSLLEGEVK 495  
RESULT 15  
GTR3\_SHEEP  
ID GTR3\_SHEEP STANDARD; PRT; 494 AA.  
AC P47843;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 26-FEB-2003 (Rel. 41, Last annotation update)  
DE Solute carrier family 2, facilitated glucose transporter, member 3  
DE (Glucose transporter type 3, brain).  
GN SLC2A3 OR GLUT3 OR GLUT-3.  
OS Ovis aries (Sheep).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Caprinae; Ovis.  
OX NCBI\_TaxID=9940;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Cocworth;  
RX MEDLINE=96109471; PubMed=8653093;  
RA Bennett B.L., Prosser C.G., Grigor M.R.;  
RT "Isolation of cDNAs and tissue specific expression of ovine glucose  
transporters";  
RL Biochem. Mol. Biol. Int. 37:9-16(1995).  
RN [2]  
RP SEQUENCE OF 88-248 FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=97392487; PubMed=9250701;  
RA Currie M.J., Bassett N.S., Gluckman P.D.;  
RT "Ovine glucose transporter-1 and -3: cDNA partial sequences and  
developmental gene expression in the placenta";  
RL Placenta 18:393-401(1997).  
CC -!- FUNCTION: Facilitative glucose transporter. Probably a neuronal  
glucose transporter.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: Highly expressed in brain.  
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose  
transporter subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L39214; AAC41629.1; -;  
CC EMBL; U9030; AAB49313.1; -;  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR InterPro; IPR005829; Sug\_transporter.  
DR InterPro; IPR003663; Sugar\_transpt.  
DR Pfam; PF00083; sugar\_tr; 1.  
DR PRINTS; PRO0171; SUGTRNSPORT.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
DR Transmembrane; Sugar transport; Transport; Glycoprotein;  
KW Multigene family; Polymorphism; Disease mutation; Diabetes mellitus.  
FT DOMAIN 1 24  
FT TRANSMEM 25 45  
FT DOMAIN 46 81  
FT TRANSMEM 82 102  
FT DOMAIN 103 111  
FT TRANSMEM 112 132  
FT DOMAIN 133 142  
FT TRANSMEM 143 163  
FT DOMAIN 164 171  
FT TRANSMEM 172 192  
FT DOMAIN 193 201  
FT TRANSMEM 202 222  
FT DOMAIN 223 287  
FT TRANSMEM 288 308  
FT DOMAIN 309 323  
FT TRANSMEM 324 344  
FT DOMAIN 345 353  
FT TRANSMEM 354 374  
FT DOMAIN 375 384  
FT TRANSMEM 385 405  
FT DOMAIN 406 417  
FT TRANSMEM 418 438  
FT DOMAIN 439 445  
FT TRANSMEM 446 466  
FT DOMAIN 467 509  
FT SITE 295  
FT SITE 489  
FT CARBOHYD 57  
FT VARIANT 78  
FT VARIANT 383  
FT VARIANT 385  
FT MUTAGEN 489  
FT CONFLICT 151  
FT SEQUENCE 509 AA; 54787 MW; 8E20CD97562C1BBF CRC64;

InterPro; IPR003663; Sugar\_transpt.  
Pfam; PF00083; sugar\_tr; 1.  
PRINTS; PRO0171; SUGTRNSPORT.  
TIGRFAMs; TIGR00879; SP; 1.  
PROSITE; PS00850; MFS; 1.  
PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
Transmembrane; Sugar transport; Transport; Glycoprotein;  
KW Multigene family; Polymorphism; Disease mutation; Diabetes mellitus.  
FT DOMAIN 1 24  
FT TRANSMEM 25 45  
FT DOMAIN 46 81  
FT TRANSMEM 82 102  
FT DOMAIN 103 111  
FT TRANSMEM 112 132  
FT DOMAIN 133 142  
FT TRANSMEM 143 163  
FT DOMAIN 164 171  
FT TRANSMEM 172 192  
FT DOMAIN 193 201  
FT TRANSMEM 202 222  
FT DOMAIN 223 287  
FT TRANSMEM 288 308  
FT DOMAIN 309 323  
FT TRANSMEM 324 344  
FT DOMAIN 345 353  
FT TRANSMEM 354 374  
FT DOMAIN 375 384  
FT TRANSMEM 385 405  
FT DOMAIN 406 417  
FT TRANSMEM 418 438  
FT DOMAIN 439 445  
FT TRANSMEM 446 466  
FT DOMAIN 467 509  
FT SITE 295  
FT SITE 489  
FT CARBOHYD 57  
FT VARIANT 78  
FT VARIANT 383  
FT VARIANT 385  
FT MUTAGEN 489  
FT CONFLICT 151  
FT SEQUENCE 509 AA; 54787 MW; 8E20CD97562C1BBF CRC64;  
Query Match 30.1%; Score 874; DB 1; Length 509;  
Best Local Similarity 38.7%; Pred. No. 6.2e-52;  
Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;  
57 LVASLAGAFSGSFLYGNLSVNVNAPTPIYKAFYNESWERRHG---RPIDPDTLTLLWSV 112  
24 LVLAFAVLSGLQFGNIGVINAPQKVIQSNETWLGQEGEPSSIPPGTLTLMAL 83  
113 TVGIFATGGLVGTLLIVKMGKVLGRKHTLANNNGFAISAALMACSLQAGAFEMLIVGRP 172  
84 SVAIFSVGGMISFLIGIISQWLGRKEAMLVNVLAVLGSLMGLANAASVEMLIIGRF 143  
173 IMGIDGVALSVLPYMLSEISPEKIRSLQGVTAIFICIGVTCOLLGLPEILGKSTWP 232  
144 LIGAYSGLTSGLVPMYVGEIATPHTLRGALGTNLQLAIVIGILIAQVLGLESLLGTASLP 203  
233 YLFGVIVPAVQLLSLPFLPDSRYLLLEKHEARAVKAFQIFLKGADYSQVEEVLAE 292  
204 LLLGLTLVPLALLQLVLPFCPESPRYLIITQNEGPARKSLKRLTGWADVSGVLAELKDE 263  
293 SHVQRSLYSVLELLRAPYVRWGVTVITMACYQLCGNAINAFWYNTSIFGKAGI-PPA 351  
264 KRKLERERPUSLLQLGSRTHRQPLIIIAVVLQSLQSGINAVFYISTSIFFIAGVGQPA 323



Db 62 AFISAVLMGFKLGSFEMILGRFIVGVCGTGTGFVPMVGVSEPTALRGALGTIHL 121  
 QY 208 FIGVFTGQLLGLPELLGKSTWYPLFGVIVPAAVQVQLSLPFLDPSRYLILLEXNEA 267  
 Db 122 GIWVGILIAQVGLDSIMGNEELWPLLSVIFIPALLOCVLLPFCPSRFLINNEEN 181  
 QY 268 RAVKAFOTFLGKADVSQVEVEVLAESHVQSRISLVSLVLELLRAPYVRQVQVIVTMACY 327  
 Db 182 RAKSVLKKLRGTADVTRDLQEMKEESQRMREKVKVITILEFRSAAYRQPIILIAVQLQSL 241  
 QY 328 QLCGLNAINWYTNISIFKAGIPPAKIPYVTLSTGGTETLAAVSGLVIEHLGRPL-LIG 386  
 Db 242 QLSGINAVFYSTSIPEKAGV--QQPYATIGSGIVNTAFTVSLFVVERAGRTLHLIG 299  
 QY 387 GFGLMGLFFGTLTTLTQLQHPVPVYLSVGLALIASCSGPGGIPFLITGCFQOQS 446  
 Db 300 LAGVAGCAV-LMTIALALLSOLPMSYLSVAIFGVAFVEVGGPFPWIVAEVLSQGP 358  
 QY 447 RPAAFIAGTWLNSFAVGLLPFFIOKSLDLYCFLVFATICTITGAILYLYFVLPETKNT 506  
 Db 359 RPAALVAGSNWTSNIVGMCFYVQVQLCGPYVFIITVLLVLFIFTYFKVPEIKGRT 418  
 QY 507 YAEISQAF-----SKRNKAYPPEE 525  
 Db 419 FDEIASGFRQGGASQDKT--PEE 440  
 RESULT 14  
 GTR4 HUMAN STANDARD; PRT; 509 AA.  
 ID GTR4 HUMAN STANDARD; PRT; 509 AA.  
 AC P14672;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Solute carrier family 2, facilitated glucose transporter, member 4  
 DE (Glucose transporter type 4, insulin-responsive).  
 DE SLC2A4 OR GLUT4.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP MEDLINE=89255193; PubMed=2856669;  
 RA Fukumoto H., Kayano T., Buse J.B., Edwards Y., Pilch P.F., Bell G.I.,  
 RA Seino S.;  
 RT "Cloning and characterization of the major insulin-responsive glucose  
 RT transporter expressed in human skeletal muscle and other insulin-  
 RL J. Biol. Chem. 264:7776-7779 (1989).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANT THR-385.  
 RX MEDLINE=93012518; PubMed=1397719;  
 RA Buse J.B., Yasuda K., Ley T.P., Seo T.S., Olson A.L., Pessin J.E.,  
 RA Karam J.H., Seino S., Bell G.I.;  
 RT "Human GLUT4/muscle-fat glucose-transporter gene. Characterization  
 RT and genetic variation."  
 RL Diabetes 41:1436-1445 (1992).  
 RN [3]  
 RP SEQUENCE OF 1-11 FROM N.A.  
 RX MEDLINE=93366193; PubMed=7916714;  
 RA Chiaramonte R., Martini R., Taramelli R., Comi P.;  
 RT "Identification of the 5' end of the gene encoding a human insulin-  
 RT responsive glucose transporter."  
 RL Gene 130:307-308 (1993).  
 RN [4]  
 RP INTERACTION WITH DAXX, AND SUMOYLATION.  
 RX MEDLINE=22028023; PubMed=11842083;  
 RA Lallot V.S., Vergara-Jauregui S., Pulido D., Sandoval I.V.;  
 RT "The insulin-sensitive glucose transporter, GLUT4, interacts  
 RT physically with Daxx. Two proteins with capacity to bind Ubc9 and  
 RT conjugated to SUMO1."  
 RL J. Biol. Chem. 277:19783-19791 (2002).

RN SUBCELLULAR LOCATION, AND MUTAGENESIS OF 489-LEU-LEU-490.  
 RP MEDLINE=94131986; PubMed=8300557;  
 RA Verhey K.J., Birnbaum M.J.;  
 RT "A Leu-Leu sequence is essential for COOH-terminal targeting signal of  
 RT GLUT4 glucose transporter in fibroblasts."  
 RL J. Biol. Chem. 269:2353-2356 (1994).  
 RN [6]  
 RP VARIANT NIDDM ILE-383.  
 RX MEDLINE=920112233; PubMed=1918382;  
 RA Kusari J., Verma U.S., Buse J.B., Henry R.R., Olefsky J.M.;  
 RT "Analysis of the gene sequences of the insulin receptor and the  
 RT insulin-sensitive glucose transporter (GLUT-4) in patients with  
 RT common-type non-insulin-dependent diabetes mellitus."  
 RL J. Clin. Invest. 88:1323-1330 (1991).  
 RN [7]  
 RP VARIANT NIDDM ILE-383.  
 RX MEDLINE=92097867; PubMed=1756912;  
 RA Choi W.H., O'Rahilly S., Buse J.B., Rees A., Morgan R., Flier J.S.,  
 RA Moller D.E.;  
 RT "Molecular scanning of insulin-responsive glucose transporter (GLUT4)  
 RT gene in NIDDM subjects."  
 RL Diabetes 40:1712-1718 (1991).  
 RN [8]  
 RP VARIANT NIDDM ILE-383.  
 RX MEDLINE=92394434; PubMed=1521731;  
 RA O'Rahilly S., Krook A., Morgan R., Rees A., Flier J.S., Moller D.E.;  
 RT "Insulin receptor and insulin-responsive glucose transporter (GLUT 4)  
 RT mutations and polymorphisms in a Welsh type 2 (non-insulin-dependent)  
 RT diabetic population."  
 RL Diabetologia 35:486-489 (1992).  
 CC -I- FUNCTION: Insulin-regulated facilitative glucose transporter.  
 CC -I- SUBUNIT: Binds to DAXX.  
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Localizes  
 CC primarily to the perinuclear region, undergoing continued  
 CC recycling to the plasma membrane where it is rapidly  
 CC reinternalized. The dileucine internalization motif is critical  
 CC for intracellular sequestration.  
 CC -I- TISSUE SPECIFICITY: Skeletal and cardiac muscles; brown and white  
 CC fat.  
 CC -I- PTM: Sumoylated.  
 CC -I- DISEASE: Defects in SLC2A4 could be a cause of noninsulin-  
 CC dependent diabetes mellitus (NIDDM) [MIM:125853]. Defects in  
 CC SLC2A4 may be a cause of certain post-receptor defects in NIDDM.  
 CC The variant in position Ile-383 is found in a small number of  
 CC NIDDM patients, but seems not to be found in nondiabetic subjects.  
 CC -I- SIMILARITY: Belongs to the sugar transporter family. Glucose  
 CC transporter subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M20747; AAA59189.1; -;  
 CC EMBL; M91463; AAA52569.1; -;  
 CC EMBL; X58489; CAA41399.1; -;  
 CC PIR; A49158; A33801.  
 CC Genew; HGNC:11009; SLC2A4.  
 CC GX; P14672; -;  
 CC MIM; 138190; -;  
 CC MIM; 125853; -;  
 CC GO; GO:0005887; C:integral to plasma membrane; TAS.  
 CC GO; GO:0005624; C:membrane fraction; TAS.  
 CC GO; GO:0005355; F:glucose transporter activity; TAS.  
 CC GO; GO:0005975; P:carbohydrate metabolism; TAS.  
 CC GO; GO:0015758; P:glucose transport; TAS.  
 CC InterPro; IPR007114; MFS.  
 CC InterPro; IPR005828; Sub transporter.  
 CC InterPro; IPR005829; Sug\_transporter.

[illegible]

RESULT 13

GTR1\_PIG

ID\_GTR1

AC P203

DT 01-E

DT	01-E
----	------

DT 28-87  
DE 0010

AD  
AD

GN SLC2

50  
875

OC Euka

OC Mamm

OX NCBI 1-1

RN [T] CEOT

RP  
SEQU

RC  
EX

RA  
Weil

```
FT TRANSMEM 186 206 6 (POTENTIAL).
FT DOMAIN 207 271
FT TRANSMEM 272 292
FT TRANSMEM 292 292
FT TRANSMEM 293 307
FT TRANSMEM 308 328
FT TRANSMEM 329 337
FT TRANSMEM 338 358
FT TRANSMEM 359 371
FT TRANSMEM 372 392
FT TRANSMEM 393 401
FT TRANSMEM 402 422
FT TRANSMEM 423 429
FT TRANSMEM 430 450
FT TRANSMEM 451 492
FT TRANSMEM 492 492
FT SITE 279 281
FT CARBOHYD 45 45
FT CONFLICT 52 52
FT CONFLICT 193 195
FT CONFLICT 357 360
FT CONFLICT 403 403
FT SEQUENCE 492 AA; 53934 MW; 4210646526F3735 CRC64;

Query Match 32.3%; Score 936.5; DB 1; Length 492;
Best Local Similarity 39.8%; Pred. No. 3.5e-56;
Matches 189; Conservative 96; Mismatches 179; Indels 11; Gaps 5;

QY 57 LVASLAGAGSSFLYGNLSVNVNAPTPYKAFYNESWERRGRPIDPTLTLLMSVTVSI 116
Db 12 LMLAVGAVLSGLQFGYNTGVNAPQKVFIEFYNTWNRHIGRIBPSTLTLLMSLSVAI 71
QY 117 FAIGLVGTLLVMKIGKVLGRKHTLLANGFAISALLMACSLQAGAFEMILVGRFNGI 176
Db 72 FSVGGMIGSPSVGLFVNRFGRNSMLMNLAFVAAVLMSFKLGSFEMILGRFTIGV 131
QY 177 DGVVALSVLPMVLSISPEKIRGSLQVTAFCIGVFTGQLLGLPELLGKESWYPLFG 236
Db 132 YCGLTGTGFVPMVGVSVSTALGALGTLHQLGIVVGLIAQVGLDSIMGNADLWPLLS 191
QY 237 VIVPAVQQLSLPLPSRBYLLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQ 296
Db 192 VIFIFALQCLLLPCSPSPRFLINNEHRAKSVLKLRTGADVFDLQEMKEEQM 251
QY 297 RSIRLVSVLELLRAPYVQVVTIVTMACYQLCLNIAWFTVNSIFKAGIPPAKIPYV 356
Db 252 MREKVTILELFRSAYQPIIAVLQSLQSLGNAVFTYSTIFEKAGV--QPVYA 309
QY 357 TLSTGQIETLAAYFSGLVIEHLGRPL-LIGFGMLGFFGTLITLLQDHAPVYLS 415
Db 310 TIGSGIVNTAFTVWSLFFVVERAGRRTLHLGLAGMAGCAV-LMTTIALALLERLPMSYLS 368
QY 416 IVGILAITASFCGPGGIPFILTGEFFQSQRPRAFIITAGTVNLSNFAVGLLPFIQKS 475
Db 369 IVAIFGVAFVFFVGPPIFWFVAELFSQGPRAIAVAGFSNWTSNFVGMCFQYEQ 429
QY 476 LTYCFLVPATCITGAIYLVFLVLPETKURTVAEISQAF-----SKRNKAVPPEE 525
Db 429 CGPYVFIITVLLVLFIFTYFKVETKGRTFDEIASGFRQGASQSDKT--PEZ 481

RESULT 12
GRL_RAT
ID GRL_RAT STANDARD; PRT; 492 AA.
AC P11167;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 1
DE (Glucose transporter type 1, erythrocyte/brain).
GN SLC2A1 OR GLUT-1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
```

FT CONFLICT 152 152 /FTID=VAR 013286.  
SQ SEQUENCE 492 AA; 54117 MW; 3B3472479834429D CRC64;  
Query Match 32.3%; Score 936.5; DB 1; Length 492;  
Best Local Similarity 39.6%; Pred. No. 3 5e-56;  
Matches 188; Conservative 98; Mismatches 178; Indels 11; Gaps 5;  
QY 57 LVASLAGAFSGSFLGYNSVNVNAPPTKAFYNESWERRHORIPDPTLTLLWSVTYSI 116  
DB 12 LMLAVGAVLSGLQFGYNTGVINAPQKVIEEFYNTQWVHRYGESILPTLTLLWSLSVAI 71  
QY 117 FAIGLIVGLTVKMGKVLGRKHTLLANNGFAISALLMACSLQAGAPEMLVGVFIMGI 176  
DB 72 FSVGGMIGSVGLFVNRFRGRNSMLMMLLAFVSAVLMGFKLQKSPSEMILGLRFFIIV 131  
QY 177 DGGVALSVLPWYLSSETSPKEISGLQGVTAIFICIGVFTGQLLGLPELLEKSTWPLYFG 236  
DB 132 YCGLTTFGVYVGEVSPFAFRGALGTQLHQLGVGLIAQVGLDSINGKDLNPLLS 191  
QY 237 VIVPAAVQVLLSPFLPDSRYLLEKNEABAVKAFOTFLGKADVSQVEEVLAESHVQ 296  
DB 192 IIFIPALLQCVLPFCFSPRELLINNEENRAKSVLKLRTADVTHDLQEMKEESQM 251  
QY 297 RSIRLSVLELRAPYRVQWVTVVIMACYQLCGLNAIWFYNSIFGHAGIPPAKIPV 356  
DB 252 MREKVTILELPSPAYROPILIAVVLQSLQSGINAVFYSTISFERAGV--QOPVYA 309  
QY 357 TLSTGTGIETLAAVFSGLVIEHLGRRL-ILGFGMLGFLPFGTILTLTLDHAPWPVYLS 415  
DB 310 TIGSGIVNTAFTVSVLFFVVERAGRTLHGLAGNAGCAI-LMTIALALLEQLPMMVYLS 368  
QY 416 IVGILAIASFSGPGGIPILTGFEPFQSQRPAAFIAGTVNLSNFAVGLLFFPIQKS 475  
DB 369 IVAIFGVFAFFVGPPIFWFVVAELFQOPRAATAVAGFNSWTSNFIWGMCFQVVEQL 428  
QY 476 LDTYCFVLVATICITGAILYLYFVLPETKRTYAEISQAF-----SKRNKAYPPEE 525  
DB 429 CGPVIIFITVLVLFYFITYFKVPTKGTPTDEIASGRQGSQSDKT--PEE 481  
RESULT 11  
GTR1\_MOUSE  
ID GTR1\_MOUSE STANDARD; PRT; 492 AA.  
AC P17809; Q61608;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Solute carrier family 2, facilitated glucose transporter, member 1  
DE (Glucose transporter type 1, erythrocyte/brain) (Glt1).  
DE SLC2A1 OR GLUT1 OR GLUT-1.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
[1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Brain;  
RC MEDLINE=89240694; PubMed=2654938;  
RX Kaesner K.H., Christy R.J., McEnithan J.C., Braiterman L.T.,  
RA Cornelius P., Pekala P.H., Lane M.D.;  
RT "Sequence, tissue distribution, and differential expression of mRNA  
for a putative insulin-responsive glucose transporter in mouse 3T3-L1  
adipocytes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 86:3150-3154 (1989).  
[2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=90274408; PubMed=2190533;  
RA Reed B.C., Shade D., Alperovich F., Vang M.;  
RT "3T3-L1 adipocyte glucose transporter (HepG2 class): sequence and  
regulation of protein and mRNA expression by insulin,  
RT differentiation, and glucose starvation.";  
RL Arch. Biochem. Biophys. 279:261-274 (1990).

[3]  
RN SEQUENCE OF 151-237 FROM N.A., AND DEVELOPMENTAL STAGE.  
RP STRAIN=C57BL/6; TISSUE=Embryo;  
RX MEDLINE=93170163; PubMed=1289053;  
RA Smith D.E., Gridley T.;  
RT "Differential screening of a PCR-generated mouse embryo cDNA library:  
glucose transporters are differentially expressed in early  
postimplantation mouse embryos.";  
RL Development 116:555-561 (1992).  
[4]  
RN SEQUENCE OF 357-463 FROM N.A., AND DEVELOPMENTAL STAGE.  
RP STRAIN=CD-1; TISSUE=Embryo;  
RX MEDLINE=92111400; PubMed=1765007;  
RA Hogan A., Heyner S., Charron M.J., Copeland N.G., Gilbert D.J.,  
RT Jenkins N.A., Thorens B., Schultz G.A.;  
RL "Glucose transporter gene expression in early mouse embryos.";  
CC Development 113:363-372 (1991).  
CC -!- FUNCTION: Facilitative glucose transporter. This isoform may be  
responsible for constitutive or basal glucose uptake. Has a very  
broad substrate specificity; can transport a wide range of aldoses  
including both pentoses and hexoses.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes  
primarily at the cell surface (by similarity).  
CC -!- DEVELOPMENTAL SPAGE: Levels decline 3-fold between days 7.5 and  
12.5 of gestation. At 7.5 dpc, expressed more strongly in  
extraembryonic tissues than in the embryo proper. Expressed in  
amion, chorion, and ectoplacental cone. In the yolk sac,  
expressed more strongly in the mesoderm layer than the ectoderm.  
CC Expression fairly widespread in the embryo at 8.5 dpc, but by 10.5  
dpc, expression is down-regulated and observed in the eye and the  
spinal cord.  
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose  
transporter subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M23384; AAA37752.1; -;  
DR EMBL; M22998; AAA37707.1; -;  
DR EMBL; X62997; CAA49367.1; -;  
DR EMBL; S77924; AAB20846.2; -;  
DR PIR; S09705; S09705.  
DR MGI; MGI:95755; SLC2a1.  
DR GO; GO:0005622; C:intracellular; IDA.  
DR GO; GO:0005886; C:plasma membrane; IDA.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR InterPro; IPR005829; Sug\_transporter.  
DR InterPro; IPR003663; Sugar\_transpt.  
DR Pfam; PF00083; sugar tr; 1.  
DR PRINTS; PR00171; SUGTRANSPT.  
DR TIGRfam; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Transmembrane; Sugar transport; Transp; Glycoprotein;  
KW Multigene family.  
FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 13 33 1 (POTENTIAL).  
FT DOMAIN 34 66 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 67 87 2 (POTENTIAL).  
FT DOMAIN 88 95 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 96 116 3 (POTENTIAL).  
FT DOMAIN 117 126 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 127 147 4 (POTENTIAL).  
FT DOMAIN 148 155 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 156 176 5 (POTENTIAL).  
FT DOMAIN 177 185 EXTRACELLULAR (POTENTIAL).

Yu W., Gibbs R.A.;  
Submitted (JUN-1998) to the EMBL/GenBank/DDJ databases.  
[4]  
SEQUENCE OF 294-423 FROM N.A.  
TISSUE=Articular cartilage;  
Nesma G., Richardson S., Bell S., Carter S., Mobasheri A.;  
"Molecular characterization and cloning of glucose transporters in  
human articular chondrocytes";  
Submitted (MAY-2001) to the EMBL/GenBank/DDJ databases.  
[5]  
VARIANT GLUT1 DEFICIENCY ILE-310.  
MEDLINE=99242249; PubMed=10227690;  
Klepper J., Wang D., Fischbarg J., Vera J.C., Jarjour I.T.,  
O'Driscoll K.R., De Vivo D.C.;  
"Defective glucose transport across brain tissue barriers: a newly  
recognized neurological syndrome";  
Neurochem. Res. 24:587-594(1999).  
[6]  
VARIANTS GLUT1 DEFICIENCY PHE-66; LEU-126; LYS-146; GLU-256 AND  
TRP-333.  
MEDLINE=20438345; PubMed=10980529;  
Wang D., Kranz-Eble P., De Vivo D.C.;  
"Mutational analysis of GLUT1 (SLC2A1) in Glut-1 deficiency  
syndrome";  
Hum. Mutat. 16:224-231(2000).  
[7]  
ERRATUM.  
Wang D., Kranz-Eble P., De Vivo D.C.;  
Hum. Mutat. 16:527-527(2000).  
[8]  
VARIANT GLUT1 DEFICIENCY HIS-126.  
MEDLINE=21487108; PubMed=11603379;  
Brockmann K., Wang D., Korenke C.G., von Moers A., Ho Y.-Y.,  
Pascual J.M., Kuang K., Yang H., Ma L., Kranz-Eble P., Fischbarg J.,  
Hanefeld F., De Vivo D.C.;  
"Autosomal dominant Glut-1 deficiency syndrome and familial  
epilepsy";  
Ann. Neurol. 50:476-485(2001).  
[9]  
VARIANT GLUT1 DEFICIENCY ASP-91.  
MEDLINE=21065632; PubMed=1136715;  
Klepper J., Willemsen M., Verrips A., Guertsen E., Herrmann R.,  
Kutzick C., Floercken A., Voit T.;  
"Autosomal dominant transmission of GLUT1 deficiency";  
Hum. Mol. Genet. 10:63-68(2001).  
-|- FUNCTION: Facilitative glucose transporter. This isoform may be  
responsible for constitutive or basal glucose uptake. Has a very  
broad substrate specificity, can transport a wide range of aldoses  
including both pentoses and hexoses.  
-|- SUBCELLULAR LOCATION: Integral membrane protein. Localizes  
primarily at the cell surface (By similarity).  
-|- TISSUE SPECIFICITY: Expressed at variable levels in many human  
tissues.  
-|- DISEASE: Defects in SLC2A1 are the cause of autosomal dominant  
GLUT1 deficiency syndrome [MIM:606777]. This disease causes a  
defect in glucose transport across the blood-brain barrier. It is  
characterized by infantile seizures, delayed development, and  
acquired microcephaly.  
-|- SIMILARITY: Belongs to the sugar transporter family. Glucose  
transporter subfamily.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; K03195; AAA52571.1; -  
EMBL; M20653; AAB61084.1; -  
EMBL; AF070544; AAC28635.1; -  
EMBL; AY034633; AAK56795.1; -

DE Solute carrier family 2, facilitated glucose transporter, member 1  
DE (Glucose transporter type 1, erythrocyte/brain).  
GN SLC2A1 OR GLUT1  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RY [1]  
RP SEQUENCE FROM N.A.  
RX TISSUE=Brain;  
RX MEDLINE=88309104; PubMed=3408493;  
RA Asano T., Shibasaki Y., Kasuga M., Kanazawa Y., Takaku F.,  
RA Akamura Y., Oka Y.;  
RT "Cloning of a rabbit brain glucose transporter cDNA and alteration of  
RT glucose transporter mRNA during tissue development.";  
RL Biochem. Biophys. Res. Commun. 154:1204-1211(1988).  
CC -!- FUNCTION: Facilitative glucose transporter. This isoform may be  
CC responsible for constitutive or basal glucose uptake. Has a very  
CC broad substrate specificity; can transport a wide range of aldoses  
CC including both pentoses and hexoses.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes  
CC primarily at the cell surface (By similarity).  
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose  
CC transporter subfamily.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M21747; AAA31444.1; -;  
DR PIR; A30797; A30797.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR InterPro; IPR005829; Sug\_transporter.  
DR InterPro; IPR003663; Sugar\_transp.  
DR Pfam; PF00083; sugar\_tr; 1  
DR PRINTS; PR00171; SUGTRNSPORT.  
DR TIGRFAMs; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Transmembrane; Sugar transport; Transport; Glycoprotein;  
KW Multigene family.  
FT DOMAIN 1 12  
FT TRANSMEM 13 33  
FT DOMAIN 34 66  
FT TRANSMEM 67 87  
FT TRANSMEM 88 95  
FT DOMAIN 96 116  
FT TRANSMEM 117 126  
FT DOMAIN 127 147  
FT TRANSMEM 148 155  
FT DOMAIN 156 176  
FT TRANSMEM 177 185  
FT DOMAIN 186 206  
FT TRANSMEM 207 271  
FT DOMAIN 272 292  
FT TRANSMEM 293 305  
FT TRANSMEM 307 327  
FT DOMAIN 328 337  
FT TRANSMEM 338 358  
FT DOMAIN 359 369  
FT TRANSMEM 370 390  
FT DOMAIN 391 401  
FT TRANSMEM 402 422  
FT DOMAIN 423 429  
FT TRANSMEM 430 450  
FT DOMAIN 451 492  
FT SITE 279

FT CARBOHYD 45 45  
SQ SEQUENCE 492 AA; 54097 MW; 575F8FC947F2B869 CRC64;  
SIMILARITY)  
Query Match 32.4%; Score 940.5; DB 1; Length 492;  
Best Local Similarity 39.7%; Pred. No. 1.9e-56;  
Matches 188; Conservative 98; Mismatches 179; Indels 9; Gaps 3;  
QY 57 LVASLAGAGSSFLYGYNLSVWNPATPYIKAFYNESWERRHGRPIDDTTLTWSVTISI 116  
DB 12 LMLAVGGVGLSLQFGYNTGVINAPQKQVIEFYNTQWHRVGERILPTTLTSLVAI 71  
QY 117 PAIGGLVGTIVRMKIGLVGRKHTLLANNGFALSALLMACSLQAGAFEMLIIVGRPMGI 176  
DB 72 PSVGGMIGSFSVGLFVNRFGRNRNMLMNLAFVSAVLMGFSKLAKSFEMLLIGRFTIGV 131  
QY 177 DGGVALSVLPMYLSSEISPKIRSLGQVTAIFICIGVFTGQLLGLPELLGKESWPLYFG 236  
DB 132 YCGLTGTFVPMYGVESFTALRGALGTIHLQIGIVVGLLIAQVFLGDSIMGNEDLWPLLS 191  
QY 237 VIVPVAVVQLSLPFLPDSRYLLEKHEARAKVAFOTFLGKADVSQEEVEVLAESHVQ 296  
DB 192 VIFPALQCIIVLPCLPESFRLINNEENRAKSVLKLRGNADVTRDLQEMKEERQM 251  
QY 297 RSIRLSVLELLRAPYVRQVTVIVTMACYQLCLNAINFYTNISIFGKAGIPPAKIPYV 356  
DB 252 MREKVTILELFRSPAYRQPILSAVVLQSLQSGINAVFYSTSIPEKAGV--QPVYA 309  
QY 357 TLTSGGIETLAAPVSGVIEHLGRPLLIQSGLMGLPFGTLITLTLQDHAPWVYLSI 416  
DB 310 TIGSGIVNTAFTVWSLFWERAGRRTHLIGLAGMAACAVLMTIALALLEQLPMSYLSI 369  
QY 417 VGLIAITASFCSGGGIPFILTGEFFQSQRPAPFAFIAGTVNLSNFAVGLLEFFIOKSL 476  
DB 370 VAIFGFVAFFVGGPIPFVIAELFSQGRPAVAVAVAGFSNWTNSFNVGCMFOYBQLC 429  
QY 477 DTGCLPVPATICITGAIYLYFVLPETKNRTVAEISQAF----SKRNKAVPPEE 525  
DB 430 GPYVFIIFTVLVLFVFFITVFKVETKGRIPDEIASGFRQGCASQSDKT--PEE 481  
RESULT 10  
GRLI\_HUMAN  
ID GRLI\_HUMAN STANDARD; PRT; 492 AA.  
AC P11166; O75535;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-JUL-1989 (Rel. 11, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Solute carrier family 2, facilitated glucose transporter, member 1  
DE (Glucose transporter type 1, erythrocyte/brain) (HepG2 glucose  
DE transporter).  
GN SLC2A1 OR GLUT1.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=85272595; PubMed=3839598;  
RA Mueckler M., Caruso C., Baldwin S.A., Panico M., Blench I.,  
RA Fukumoto H., Seino S., Imura H., Lienhard G.B., Lodish H.F.;  
RT "Sequence and structure of a human glucose transporter.";  
RL Science 229:941-945(1985).  
RN [2]  
RP SEQUENCE OF 1-6 FROM N.A.  
RX MEDLINE=88196648; PubMed=2834252;  
RA Fukumoto H., Seino S., Imura H., Seino Y., Bell G.I.;  
RT "Characterization and expression of human HepG2/erythrocyte glucose-  
RT transporter gene.";  
RL Diabetes 37:657-661(1988).  
RN [3]  
RP SEQUENCE OF 150-492 FROM N.A.  
RC TISSUE=Brain;

Db 191 FIFVPALLOCIILPFAPESEPRFLLNRNEENKAKSVLKKLGGTTDVSSDLQEWKSESRQM 250  
QY 297 RSRLSVLELLRAPVVRWQVTVITMACYQLCGLNAIWFYTNIGFGKAGIPPAKIPYV 356  
Db 251 MREKKTIVLELFRSPMYRQPIIAIVLQLSQGLSGLNAVYFSTIFERKSGV--EQPVYA 308  
QY 357 TLSTGGIETLAAVFGSLVIEHGLRRPL-LIGFGLMGLFFGLTITLTLQDHAPWVPYLS 415  
Db 309 TIGSGVNTAFTVSVLFWVERAGRTLHLIGLAGMAGCAI-LMTIALTLDDQWPMWSYLS 367  
QY 416 IVCILAIISFSCGPGGIPFILTGEFFQSQORPAAFIAGTVNLSNFAVGLLEPFIQKS 475  
Db 368 IVAIFGVAFPEIFGPGIPWFIWAELEFSGQPRPAFAVAGLSNWTSNFIVGMPQIAOL 427  
QY 476 LDYFCILVPATICITGAIYLYFVLPTKNTYAEISQAF-----SKRNKAYPEE 525  
Db 428 CGSIVFIIFTVLLVLFIFFTYFKVPETKGTDFEIAVRRQGGASQDKT--PDE 480

RESULT 8

GTR3\_CHICK  
ID GTR3\_CHICK STANDARD; PRT; 496 AA.  
AC P28568;  
DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-DEC-1992 (Rel. 24, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Solute carrier family 2, facilitated glucose transporter, member 3  
DE (Glucose transporter type 3) (CER-GT3).  
GN SLC2A3 OR GLUT3.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauia; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91342646; PubMed=1875932;  
RA White M.K.; Rall T.B.; Weber M.J.;  
RT "Differential regulation of glucose transporter isoforms by the arc  
RT oncogene in chicken embryo fibroblasts."  
RL Mol. Cell. Biol. 11:4448-4454(1991).  
CC -!- FUNCTION: Facilitative glucose transporter. Probably a neuronal  
CC glucose transporter.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose  
CC transporter subfamily.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M37785; AAA48662.1; -  
CC InterPro; IPR007114; MFS.  
CC InterPro; IPR005828; Sub\_transporter.  
CC InterPro; IPR005829; Sug\_transporter.  
CC InterPro; IPR003663; Sugar\_transpt.  
CC Pfam; PF00083; sugar\_tr; 1.  
CC PRINTS; PR00171; SUGTRNSPORT.  
CC TIGRFAMs; TIGR00879; SP; 1.  
CC PROSITE; PS50850; MFS; 1.  
CC PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
CC PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
CC Transmembrane; Sugar transporter; Transport; Glycoprotein;  
KW Multigene family.  
KW DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 7 27 1 (POTENTIAL).  
FT DOMAIN 28 65 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 66 86 2 (POTENTIAL).  
FT DOMAIN 87 94 CYTOPLASMIC (POTENTIAL).  
FT

FT TRANSMEM 95 115 3 (POTENTIAL).  
FT DOMAIN 116 125 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 126 146 4 (POTENTIAL).  
FT DOMAIN 147 154 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 155 175 5 (POTENTIAL).  
FT DOMAIN 176 184 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 185 205 6 (POTENTIAL).  
FT DOMAIN 206 270 7 (POTENTIAL).  
FT TRANSMEM 271 291 8 (POTENTIAL).  
FT DOMAIN 292 306 9 (POTENTIAL).  
FT TRANSMEM 307 327 10 (POTENTIAL).  
FT DOMAIN 328 335 11 (POTENTIAL).  
FT TRANSMEM 336 356 12 (POTENTIAL).  
FT DOMAIN 357 367 10 (POTENTIAL).  
FT TRANSMEM 368 388 11 (POTENTIAL).  
FT DOMAIN 389 398 12 (POTENTIAL).  
FT TRANSMEM 399 419 12 (POTENTIAL).  
FT DOMAIN 420 426 12 (POTENTIAL).  
FT TRANSMEM 427 447 12 (POTENTIAL).  
FT DOMAIN 448 496 12 (POTENTIAL).  
FT SITE 278 280  
FT CARBOHYD 44 44  
FT SEQUENCE 496 AA; 54174 MW; 75B3C0F61A7A92A5 CRC64;  
SQ  
Query Match 32.5%; Score 942.5; DB 1; Length 496;  
Best Local Similarity 39.3%; Pred. No. 1.4e-56;  
Matches 193; Conservative 94; Mismatches 187; Indels 17; Gaps 5;  
QY 49 RDKWSCSLVSLAGAFSGSSFLYGVNLVSVNAPPYTKAFVNESWERRHGRIPDPDTLTL 108  
Db 4 KKKITASLIYAVSAAIG-SLQFGYNTGVINAPKIIQAFYNRTLSORSGTISPELLTS 62  
QY 109 LWSVTVSIFAAGLVGTLIVKMGVGRKHTLANNNGFALSAALLMACSIQAQAFEMLI 168  
Db 63 LWSVSAIFSGVMIGSVSVSLFFNRFGRNSMLLVNVLAPAGGALMALSIAKAVEMLI 122  
QY 169 VGRFTMGIDGGVALSVLPMYLSSEISPKIEIRGLSGQVTAIFTCIGVFTQQLGLPELLGKE 228  
Db 123 IGRFTIIGLFCGLCTGCFVPMYISEVPTSLRGAFGLNQLGIWGLVVAIQIGLEGIMGTE 182  
QY 229 STWPLYFGVIVPAVQVLLSPFLPDSPRYLLEKHNBARAKAFQTFGLKADYSQVEE 288  
Db 183 ALWPLLGLFTIIPAVLQCVALLFCPESPRFLINKWEEKCAQTVLQKLRGTQDVSDISE 242  
QY 289 VLAESHVQSRISLNSVLELLRAPYVRWQVTVITMACYQLCGLNAIWFYTNISIFGKAGI 348  
Db 243 MKEESAKWSQEKKAIVLELFRSPNVRQPIIISITLQLSQGLSGLNAVYFSTIFERAGI 302  
QY 349 PPAKIPYVTLSTGGIETLAAVFGSLVIEHGLRRPILLIGFGMLGLFFGTTLTITLLODHA 408  
Db 303 --TQPVYATIGAGVNTVTVVSLFLVERAGRTLHIVLGMVCAAVMTIALAKE-- 358  
QY 409 PWVPYLSIVGILAIISAFSCGPGGIPFILTGEFFQSQORPAAFIAGTVNLSNFAVGLL 468  
Db 359 KWRIVISIVATFGFVALFEIGPGIPWFIWAELEFSGQPRPAFAVAGCSNWTSNFLVGM 418  
QY 469 PFFIOKSLDTCFLVPATICITGAIYLYFVLPTKNTYAEISQAFSKRNKAYPEEKID 528  
Db 419 FPYAEKLCGPYVFLIFLVLFFLIFFITFTYFKVPETKGTDFEDISRGF-----EQVE 469  
QY 529 SAVTDAPASSP 539  
Db 470 ---TSPPSSPP 477  
RESULT 9  
GTRL\_RABIT  
ID GTRL\_RABIT STANDARD; PRT; 492 AA.  
AC P13355;  
DT 01-JAN-1990 (Rel. 13, Created)  
DT 01-JAN-1990 (Rel. 13, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DT



primarily at the cell surface (By similarity).

-1- SIMILARITY: Belongs to the sugar transporter family. Glucose transporter subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; L07300; AAB02037.1; --  
 InterPro; IPR007114; MFS.  
 InterPro; IPR005828; Sub transporter.  
 InterPro; IPR005829; Sugar transporter.  
 InterPro; IPR003663; Sugar transp.  
 Pfam; PF00083; sugar tr. 1.  
 PRINTS; PR00171; SUGRTRANSPORT.  
 TIGRFAMs; TIGR00879; SP; 1.  
 PROSITE; PS00850; MFS; 1.  
 PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 Transmembrane; Sugar transport; Transport; Glycoprotein;  
 Multigene family.

DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).  
 TRANSSEM 12 32 1 (POTENTIAL).  
 DOMAIN 33 65 EXTRACELLULAR (POTENTIAL).  
 TRANSSEM 66 86 2 (POTENTIAL).  
 DOMAIN 87 94 CYTOPLASMIC (POTENTIAL).  
 TRANSSEM 95 115 3 (POTENTIAL).  
 DOMAIN 116 125 EXTRACELLULAR (POTENTIAL).  
 TRANSSEM 126 146 4 (POTENTIAL).  
 DOMAIN 147 154 CYTOPLASMIC (POTENTIAL).  
 TRANSSEM 155 175 5 (POTENTIAL).  
 DOMAIN 176 184 EXTRACELLULAR (POTENTIAL).  
 TRANSSEM 185 205 6 (POTENTIAL).  
 DOMAIN 206 270 CYTOPLASMIC (POTENTIAL).  
 TRANSSEM 271 291 7 (POTENTIAL).  
 DOMAIN 292 306 EXTRACELLULAR (POTENTIAL).  
 TRANSSEM 307 327 8 (POTENTIAL).  
 DOMAIN 328 336 CYTOPLASMIC (POTENTIAL).  
 TRANSSEM 337 357 9 (POTENTIAL).  
 DOMAIN 358 369 EXTRACELLULAR (POTENTIAL).  
 TRANSSEM 370 390 10 (POTENTIAL).  
 DOMAIN 391 400 CYTOPLASMIC (POTENTIAL).  
 TRANSSEM 401 421 11 (POTENTIAL).  
 DOMAIN 422 428 EXTRACELLULAR (POTENTIAL).  
 TRANSSEM 429 449 12 (POTENTIAL).  
 DOMAIN 450 490 CYTOPLASMIC (POTENTIAL).  
 SITE 177 179 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).  
 CARBOHYD 44 44 N-LINKED (GLCNAC...) (POTENTIAL).  
 SEQUENCE 490 AA; 54086 MW; 63BB176812D06104 CRC64;

Query Match 32.6%; Score 946.5; DB 1; Length 490;  
 Best Local Similarity 40.6%; Pred. No. 7.4e-57;  
 Matches 193; Conservative 95; Mismatches 176; Indels 11; Gaps 5;

Qy 57 LVASLAGAFSSFLYGVNLSVNNAPTPIKAFYNESWERRHGRPIDPDTLLMSVTVSI 116  
 Db 11 LMLAVGAVLSGLQFGYNTGVINRPOKVEDFNTHTWLYREBEIPATLTLWSLSVAI 70  
 Qy 117 FAIGGLVGTLLIVKMGKVLGRKHTLLANNGFAISAALLMACSLQAGAFEMLIYGRFTMGI 176  
 Db 71 FSVGGMIGSFVGLFVNRFGRRNMLNLAFLAFLVLMGFSKVALSFEMLILGRFTIIGL 130  
 Qy 177 DGGVALSVLPYLSPEIKRISGLQVTAIFTCIGVFTQGLLGLPELLGKSTWPLYFG 236  
 Db 131 YSGLTGTFVPMYGVESPTALRGALGTGHQIGIVLGILIAQVFGDLIMGDSLWPLLLG 190  
 Qy 237 VIVVPAVQVLLSPLPDSRYLLLEKHNARAVKAFQTFGLKADYSQEVVEVLAEHVQ 296

DOMAIN 423 429 EXTRACELLULAR (POTENTIAL).  
 TRANSSEM 430 450 12 (POTENTIAL).  
 DOMAIN 451 492 CYTOPLASMIC (POTENTIAL).  
 SITE 279 281 DEFINES SUBSTRATE SPECIFICITY (BY SIMILARITY).  
 CARBOHYD 45 45 N-LINKED (GLCNAC...) (POTENTIAL).  
 SEQUENCE 492 AA; 54131 MW; D83605D0F13D7252 CRC64;

Query Match 32.7%; Score 949.5; DB 1; Length 492;  
 Best Local Similarity 40.0%; Pred. No. 4.6e-57;  
 Matches 190; Conservative 100; Mismatches 174; Indels 11; Gaps 5;

Qy 57 LVASLAGAFSSFLYGVNLSVNNAPTPIKAFYNESWERRHGRPIDPDTLLMSVTVSI 116  
 Db 12 LMLAVGAVLSGLQFGYNTGVINRPOKVEDFNTHTWLYREBEIPATLTLWSLSVAI 71  
 Qy 117 FAIGGLVGTLLIVKMGKVLGRKHTLLANNGFAISAALLMACSLQAGAFEMLIYGRFTMGI 176  
 Db 72 FSVGGMIGSFVGLFVNRFGRRNMLNLAFLAFLVLMGFSKVALSFEMLILGRFTIIGV 131  
 Qy 177 DGGVALSVLPYLSPEIKRISGLQVTAIFTCIGVFTQGLLGLPELLGKSTWPLYFG 236  
 Db 132 YSGLTGTFVPMYGVESPTALRGALGTGHQIGIVLGILIAQVFGDLIMGDSLWPLLLS 191  
 Qy 237 VIVVPAVQVLLSPLPDSRYLLLEKHNARAVKAFQTFGLKADYSQEVVEVLAEHVQ 296  
 Db 192 VIFIPALQOCHLLPCPSPRFLNRRNEHRAKSVLKURGTADVTRDLQEKESRQM 251  
 Qy 297 RSIRLVSVLELLRAPYVQVVTIVTWACYQLCGLNAINFYTNISIFGKAGIPPAKIPYV 356  
 Db 252 MREKVKVILELFRSAAYQPILIAVVLQSQLSGINAVFYVSISPEKAGV--QPYYA 309  
 Qy 357 TLSTGGIETLAAYFSGVLVIEHGRPL-LIGFGMLGLFFGTLLITLTLQDHAPVPLYLS 415  
 Db 310 TIGSGIVNTAFTVSVLFFVVERAGRRTLHLIGLAGMAGCAV-LMTIALALLERLPWMSYLS 368  
 Qy 416 IGVLLAIATSCSGGGIPFLITGFFGFOOSQRPAAFIAGTVNLSFAVGLLPFFIOKS 475  
 Db 369 IVAIFGFVAFVGVGPPIFWFVAFELFSGQRPAAIAVAGFSNNTSFIVMGCFQVBEQL 428  
 Qy 476 LDYVCLFVATICITGAYLVPLVETKNRTVAEISQAF----SKRNKAYPPEE 525  
 Db 429 GCPYVFIITVLLVLVFFITFKVETKGRFDEIASGFRGASQSQXT--PEE 481

RESULT 7  
 GTRI\_CHICK STANDARD; PRT; 490 AA.  
 AC P46896;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Solute carrier family 2, facilitated glucose transporter, member 1  
 DE (Glucose transporter type 1) (GTL1).  
 GN SLC2A1 OR GLUT1.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96157892; PubMed=8589457;  
 RA Wagstaff P., Kang H.Y., Mylott D., Robbins P.J., White M.K.;  
 RT "Characterization of the avian GLUT1 glucose transporter:  
 RT differential regulation of GLUT1 and GLUT3 in chicken embryo  
 RT fibroblasts";  
 RL Mol. Biol. Cell 6:1575-1589(1995).  
 CC -!- FUNCTION: Facilitative glucose transporter. This isoform may be  
 CC responsible for constitutive or basal glucose uptake. Has a very  
 CC broad substrate specificity; can transport a wide range of alcohols  
 CC including both pentoses and hexoses (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes

ID	GTRI_BOVIN	STANDARD;	PRT;	492 AA.
AC	P27674;			
AD	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Solute carrier family 2, facilitated glucose transporter, member 1			
DE	(Glucose transporter type 1, erythrocyte/brain).			
GN	SLC2A1 OR GLUT1 OR GLUT-1.			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Boado R.J., Pardridge W.M.;			
RT	"Molecular cloning of the bovine blood-brain barrier glucose			
RT	transporter cDNA and demonstration of phylogenetic conservation of the			
RT	5' untranslated region."			
RL	Mol. Cell. Neurosci. 1:224-232(1991).			
CC	-!- FUNCTION: Facilitative glucose transporter. This isoform may be			
CC	responsible for constitutive or basal glucose uptake. Has a very			
CC	broad substrate specificity; can transport a wide range of aldoses			
CC	including both pentoses and hexoses.			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein. Localizes			
CC	primarily at the cell surface (By similarity).			
CC	-!- SIMILARITY: Belongs to the sugar transporter family. Glucose			
CC	transporter subfamily.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; M60448; AAA30550.1; -			
DR	PIR; I45902; I45902. MFS.			
DR	InterPro; IPR007114; MFS.			
DR	InterPro; IPR005828; Sub transporter.			
DR	InterPro; IPR005829; Sug transporter.			
DR	InterPro; IPR003663; Sugar transp.			
DR	Pfam; PF00083; sugar tr_1			
DR	PRINTS; PRC0171; SUGTRNSPORT.			
DR	TIGRFAMs; TIGR00879; SP; 1.			
DR	PROSITE; P50850; MFS; 1.			
DR	PROSITE; P500216; SUGAR_TRANSPORT_1; 1.			
DR	PROSITE; P500217; SUGAR_TRANSPORT_2; 1.			
KW	Transmembrane; Sugar transport; Transport; Glycoprotein;			
KW	Multigene family.			
FT	DOMAIN 1 12			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 13 33			1 (POTENTIAL).
FT	DOMAIN 34 66			EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 67 87			2 (POTENTIAL).
FT	DOMAIN 88 95			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 96 116			3 (POTENTIAL).
FT	DOMAIN 117 126			EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 127 147			4 (POTENTIAL).
FT	DOMAIN 148 155			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 156 176			5 (POTENTIAL).
FT	DOMAIN 177 185			EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 186 206			6 (POTENTIAL).
FT	DOMAIN 207 271			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 272 292			7 (POTENTIAL).
FT	DOMAIN 293 306			EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 307 327			8 (POTENTIAL).
FT	DOMAIN 328 337			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 338 358			9 (POTENTIAL).
FT	DOMAIN 359 368			EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 369 389			10 (POTENTIAL).
FT	DOMAIN 390 401			CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 402 422			11 (POTENTIAL).

RESULT 6  
GTRI\_BOVIN

DR InterPro; IPR005829; Sug\_transporter.  
DR InterPro; IPR003663; Sugar\_transp.  
DR Pfam; PF00083; sugar tr; 1.  
DR PRINTS; PR00171; SUGTRNSPORT.  
DR TIGRFams; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
DR Transmembrane; Sugar transport; Glycoprotein;  
KW Multigene family.  
KW DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 12 32 1 (POTENTIAL).  
FT DOMAIN 12 32 1 (POTENTIAL).  
FT TRANSMEM 33 67 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 68 88 2 (POTENTIAL).  
FT TRANSMEM 89 100 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 101 120 3 (POTENTIAL).  
FT TRANSMEM 121 131 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 122 152 4 (POTENTIAL).  
FT TRANSMEM 153 160 5 (POTENTIAL).  
FT DOMAIN 161 181 6 (POTENTIAL).  
FT TRANSMEM 182 192 6 (POTENTIAL).  
FT DOMAIN 193 213 7 (POTENTIAL).  
FT TRANSMEM 214 276 7 (POTENTIAL).  
FT DOMAIN 277 297 8 (POTENTIAL).  
FT TRANSMEM 298 314 8 (POTENTIAL).  
FT DOMAIN 315 335 9 (POTENTIAL).  
FT TRANSMEM 336 341 9 (POTENTIAL).  
FT DOMAIN 342 362 10 (POTENTIAL).  
FT TRANSMEM 363 367 10 (POTENTIAL).  
FT DOMAIN 368 388 11 (POTENTIAL).  
FT TRANSMEM 389 411 11 (POTENTIAL).  
FT DOMAIN 412 432 12 (POTENTIAL).  
FT TRANSMEM 433 438 12 (POTENTIAL).  
FT DOMAIN 439 459 12 (POTENTIAL).  
FT TRANSMEM 460 501 12 (POTENTIAL).  
FT CARBOHYD 50 50 N-LINKED (GLCNAC... (POTENTIAL).  
FT SEQUENCE 501 AA; 55524 MW; 3696501DE11DC5C1 CRC64;  
Query Match 37.9%; Score 1099.5; DB 1; Length 501;  
Best Local Similarity 45.1%; Pred. No. 3.4e-67;  
Matches 214; Conservative 94; Mismatches 163; Indels 3; Gaps 1;  
QY 56 LLVASLAGAFGSFLYGVNLSVNVNAPTPIKAFYNESWERRRGRPIDPTLLLSVTVS 115  
DB 16 LALATLJIAAFGSFGYGVNVAVNSPSEFMQFYNDYDREENIESFTLLLSLTVS 75  
QY 116 IFAIGLVGTLLVMKIGVLGRKHTLLANNFPAISALLMACSLOQAFEMLIVGRFPG 175  
DB 76 MFPPGGFISLLVGNLVNKLGRKALLFNFIPLPAIPMGCSQIAQSPELLIISRLVVG 135  
QY 176 IDGVVALSVLPMYLSISPKETRGSLGQVTAIPICIGVTGQLGLPELLGKSTWPLYF 235  
DB 136 ICAGISSNVVPMYLGELAPKRLGALGVVPQLFITVGLVLAQLFGLRLSLANEDGNPVL 195  
QY 236 GVIVVPAVQLLSLPLPDSRVLLEKHEARAVKAFOTFLGKADVSQVEEVLAEHV 295  
DB 196 GLTGVPAGLQLLPLFPFSPRYLLIORKDDEAAERALTQIRGKDVHLEMEIRKDEBA 255  
QY 296 QRSIRLVSVLELRLAPYVQVWVTVVTVMACVQLCGNNAIFVYTNISIFGKAGIPPAKIPY 355  
DB 256 EXAAGFISVWKLFTMQLSLRWQLISMVLMAGQQLSGVNAIYYADIIYLSAGVKSDVQY 315  
QY 356 VTLSGTGIIETLAIVFSLGVIHLGRPLLIIGFGLMGLPFGTLITITLQDHAPVWYLS 415  
DB 316 VTAGTGAIVNVFMILITIFVVELWGRFRLLLVGSPSTCLIACLVLTAALALQNTISWMPYIS 375  
QY 416 IVGILAIIASFCGPGGIPILTGEFFQSQSORPAAFIAGTVNLSNFAVGLLFPQIKS 475  
DB 376 IVCVIVVIVGHALGPSIPALLITEIFLOSSRPAAYMIGSVHLSNFTVGLIFPFIQMG 435  
QY 476 LDTYCLVLPATICITGAIYLYFVLPETKNTYAEISQAFSKRNK---AYPPEEK 526  
DB 436 LGPYSFIIFATICITLTIYLYFVVPETKGTFTIELIIFIMKNKVSQDVTPKKEE 489

RESULT 5

GTLL\_HUMAN STANDARD; PRT; 496 AA.  
AC Q9BYW1;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 15-MAR-2004 (Rel. 43, Last annotation update)  
DE Solute carrier family 2, facilitated glucose transporter, member 11  
DE (Glucose transporter type 11) (Glucose transporter type 10).  
GN SLC2A11 OR GLUT11 OR GLUT10.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
SEQUENCE FROM N.A. (ISOFORM 1).  
TISSUE=Heart;  
DOEGE H., Bocianski A., Scheepers A., Axer H., Eckel J., Joost H.-G.;  
"Two novel members of the family of sugar transport facilitators:  
genomic organization and predominant expression in lymphocytes (GLUT9)  
and muscle (GLUT10).";  
Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
[2]  
SEQUENCE FROM N.A. (ISOFORM 2).  
MEDLINE=21616996; PubMed=11741323;  
RA Sasaki T., Minoshima S., Shiohama A., Shintani A., Shimizu A.,  
Asakawa S., Kawasaki K., Shimizu N.;  
"Molecular cloning of a member of the facilitative glucose  
transporter gene family, GLUT11 (SLC2A11) and identification of  
transcription variants.";  
RT Biochem. Biophys. Res. Commun. 289:1218-1224(2001).  
RL Biochem. Biophys. Res. Commun. 289:1218-1224(2001).  
CC -!- FUNCTION: Facilitative glucose transporter (By similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- ALTERNATIVE PRODUCTS.  
CC Event=Alternative splicing; Named isoforms=2;  
CC Name=1; Synonyms=GLUT11;  
CC IsoId=Q9BYW1-1; Sequence=Displayed;  
CC Name=2; Synonyms=GLUT11-s;  
CC IsoId=Q9BYW1-2; Sequence=VSP\_006293, VSP\_006294, VSP\_006295;  
CC -!- TISSUE SPECIFICITY: Expressed in heart and skeletal muscle.  
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE  
TRANSPORTERS SUBFAMILY.  
CC -!- CAUTION: Has been described as GLUT10 in literature, but this  
gene name has already been used for SLC2A10.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
EMBL; AJ271290; CAC29020.1; -;  
EMBL; AB049214; BAB68410.1; -;  
DR Genew; HGNC:14239; SLC2A11.  
DR InterPro; IPR007114; MFS.  
DR InterPro; IPR005828; Sub\_transporter.  
DR InterPro; IPR005829; Sug\_transporter.  
DR InterPro; IPR003663; Sugar\_transp.  
DR Pfam; PF00083; sugar\_tr; 1.  
DR PRINTS; PR00171; SUGTRNSPORT.  
DR TIGRFams; TIGR00879; SP; 1.  
DR PROSITE; PS00850; MFS; 1.  
DR PROSITE; PS00216; SUGAR\_TRANSPORT\_1; FALSE\_NEG.  
DR PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Transpotein; Sugar transport; Transmembrane; Glycoprotein;  
KW Multigene family; Alternative splicing.  
FT DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 12 32 1 (POTENTIAL).  
FT DOMAIN 33 67 EXTRACELLULAR (POTENTIAL).  
FT

RA Kasahara T., Kasahara M.,  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: Cytochalasin B-sensitive carrier. Seems to function  
 CC primarily as a fructose transporter.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed in small intestine, and at much  
 CC lower levels in kidney, skeletal muscle, and adipose tissue.  
 CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose  
 CC transporter subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L05195; AAA02627.1; -;  
 CC EMBL; D13871; BAA02983.1; -;  
 CC EMBL; D28562; BAA05912.1; -;  
 CC PIR; I53268; I53268.  
 CC InterPro; IPR007114; MFS.  
 CC InterPro; IPR005828; Sub\_transporter.  
 CC InterPro; IPR005829; Sug\_transporter.  
 CC InterPro; IPR003663; Sugar\_transpt.  
 CC Pfam; PF00083; sugar\_tr; 1.  
 CC PRINTS; PR00171; SUGETRNSPORT.  
 CC TIGRams; TIGR00879; SP; 1.  
 CC PROSITE; PS00850; MFS; 1.  
 CC PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
 CC PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
 CC Transmembrane; Sugar transport; Transp.; Glycoprotein;  
 CC Multigene family.  
 KW DOMAIN 1 11 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 1 11  
 FT TRANSSEM 12 32  
 FT DOMAIN 33 71  
 FT TRANSSEM 72 92  
 FT DOMAIN 93 99  
 FT TRANSSEM 100 120  
 FT DOMAIN 121 131  
 FT TRANSSEM 132 152  
 FT DOMAIN 153 160  
 FT TRANSSEM 161 181  
 FT DOMAIN 182 191  
 FT TRANSSEM 192 212  
 FT DOMAIN 213 276  
 FT TRANSSEM 277 297  
 FT DOMAIN 298 314  
 FT TRANSSEM 315 335  
 FT DOMAIN 336 342  
 FT TRANSSEM 343 363  
 FT DOMAIN 364 367  
 FT TRANSSEM 368 388  
 FT DOMAIN 389 411  
 FT TRANSSEM 412 432  
 FT DOMAIN 433 438  
 FT TRANSSEM 439 459  
 FT DOMAIN 460 502  
 FT TRANSSEM 503 523  
 FT CONFLICT 164 164 V -> A (IN REF. 1).  
 FT CONFLICT 285 285 A -> T (IN REF. 2).  
 FT CONFLICT 452 452 T -> S (IN REF. 2).  
 FT CONFLICT 479 479 K -> N (IN REF. 1).  
 SQ SEQUENCE 502 AA; 55543 MW; 871FEDEF4E27884D CRC64;

Query Match 38.3%; Score 1112.5; DB 1; Length 502;  
 Best Local Similarity 44.9%; Pred. No. 4.5e-68;  
 Matches 218; Conservative 96; Mismatches 167; Indels 5; Gaps 2;  
 QY 56 LVLASLAGAGSGFLYGNLSVNVNAPTVIKAFYNESWERRHGRIPDPTLLWSVTVS 115  
 DB 16 LALATFLAAGSGFQYGNVAVNSPBMQOQFYNDYDRNKENIESFTLTLWSLTVS 75

QY 116 IFAGGLVGTLLIVNMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMG 175  
 DB 76 MFPGFGFGLSLVMGVLVNNLGRKGLALLFNNI:FSILPAILMGCSKIAXSFEIILASRLVVG 135  
 QY 176 IDGVALSVLPWYLSEISPKIRGSLGQVTAIFCIGVFTGQLLGLLPELLOKSTWYLPF 235  
 DB 136 ICAGISSNVPMYLGELAPKNLRGALGVVPQLFITVGILVAQLFGLRSVLASEGWPILL 195  
 QY 236 GVIVVPAVVQLLSLRFLPDSPRYLLEKHNARAVKAFOTELGKADVSQEEVEVLASHV 295  
 DB 196 GUTGPDAGLQQLLLFFPESPRYLLIQKNESAELALQTLRGKVDVMEKEERKDEA 255  
 QY 296 QRSIRLVSVLELLRAPPYRWQVTVIVTMACYQLGLNAINWFTYNSIFGKAGIPAKIPY 355  
 DB 256 EKAAAGFISVWKLFRMQSLRWQLISTIVLMAGQQLSGVNAIYYADQIYLSAGVSNVDVQY 315  
 QY 356 VTLSTGGIETLAAPSGVIEHLGRPLLIIGGGLMGLFFGTLTITLTLODHPWVPYLS 415  
 DB 316 VTAGTGAVNVFMTVTVFVVELWGRNLL:GFSTCLTACIVLVALALQNTISWMPYVS 375  
 QY 416 IVGILAIIASFCSGFGGIPFILTGEFFQSQORPAFIAGTVNWSLNFVGLLPPFIQKS 475  
 DB 376 IVCVIVYVIGHAVGSPIPALFITEIFLOSSEPRAYMIGSGVHWSLNFVGLLPPFIQVG 435  
 QY 476 LQTYCLVLPATICITGAILYLVFVLPETKNRYVAISQAFSGKNGK---AYPPEEKIDSAVT 532  
 DB 436 LQPYSFIIIPAIICLLTIVFMVDETKRGTFRVEINQIFAKKNKVDVYP--EKEEKELN 493  
 QY 533 DAPASS 538  
 DB 494 DLPAT 499  
 RESULT 4  
 GTRS\_MOUSE  
 ID GTRS\_MOUSE STANDARD; PRT; 501 AA.  
 AC Q9WV38;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Solute carrier family 2, facilitated glucose transporter, member 5  
 DE (Glucose transporter type 5, small intestine) (Fructose transporter).  
 GN SLC2A5 OR GLUT5.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6;  
 RX MEDLINE=22027171; PubMed=12031501;  
 RA Kwon O., Levine M., Burant C.F.;  
 RT "Cloning and functional characterization of the mouse fructose  
 RT transporter, GLUT5."  
 RL Biochim. Biophys. Acta 1576:191-197(2002).  
 CC -!- FUNCTION: Cytochalasin B-sensitive carrier. Seems to function  
 CC primarily as a fructose transporter (By similarity).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE  
 CC TRANSPORTERS SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF161071; AAD42235.1; -;  
 CC MGD; MG1:1928369; SLC2A5.  
 CC InterPro; IPR007114; MFS.  
 CC InterPro; IPR005828; Sub\_transporter.

[illegible]

51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546  
 547  
 548  
 549  
 550  
 551  
 552  
 553  
 554  
 555  
 556  
 557  
 558  
 559  
 560  
 561  
 562  
 563  
 564  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612  
 613  
 614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629  
 630  
 631  
 632  
 633  
 634  
 635  
 636  
 637  
 638  
 639  
 640  
 641  
 642  
 643  
 644  
 645  
 646  
 647  
 648  
 649  
 650  
 651  
 652  
 653  
 654  
 655  
 656  
 657  
 658  
 659  
 660  
 661  
 662  
 663  
 664  
 665  
 666  
 667  
 668  
 669  
 670  
 671  
 672  
 673  
 674  
 675  
 676  
 677  
 678  
 679  
 680  
 681  
 682  
 683  
 684  
 685  
 686  
 687  
 688  
 689  
 690  
 691  
 692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702  
 703  
 704  
 705  
 706  
 707  
 708  
 709  
 710  
 711  
 712  
 713  
 714  
 715  
 716  
 717  
 718  
 719  
 720  
 721  
 722  
 723  
 724  
 725  
 726  
 727  
 728  
 729  
 730  
 731  
 732  
 733  
 734  
 735  
 736  
 737  
 738  
 739  
 740  
 741  
 742  
 743  
 744  
 745  
 746  
 747  
 748  
 749  
 750  
 751  
 752  
 753  
 754  
 755  
 756  
 757  
 758  
 759  
 760  
 761  
 762  
 763  
 764  
 765  
 766  
 767  
 768  
 769  
 770  
 771  
 772  
 773  
 774  
 775  
 776  
 777  
 778  
 779  
 780  
 781  
 782  
 783  
 784  
 785  
 786  
 787  
 788  
 789  
 790  
 791  
 792  
 793  
 794  
 795  
 796  
 797  
 798  
 799  
 800  
 801  
 802  
 803  
 804  
 805  
 806  
 807  
 808  
 809  
 810  
 811  
 812  
 813  
 814  
 815  
 816  
 817  
 818  
 819  
 820  
 821  
 822  
 823  
 824  
 825  
 826  
 827  
 828  
 829  
 830  
 831  
 832  
 833  
 834  
 835  
 836  
 837  
 838  
 839  
 840  
 841  
 842  
 843  
 844  
 845  
 846  
 847  
 848  
 849  
 850  
 851  
 852  
 853  
 854  
 855  
 856  
 857  
 858  
 859  
 860  
 861  
 862  
 863  
 864  
 865  
 866  
 867  
 868  
 869  
 870  
 871  
 872  
 873  
 874  
 875  
 876  
 877  
 878  
 879  
 880  
 881  
 882  
 883  
 884  
 885  
 886  
 887  
 888  
 889  
 890  
 891  
 892  
 893  
 894  
 895  
 896  
 897  
 898  
 899  
 900  
 901  
 902  
 903  
 904  
 905  
 906  
 907  
 908  
 909  
 910  
 911  
 912  
 913  
 914  
 915  
 916  
 917  
 918  
 919  
 920  
 921  
 922  
 923  
 924  
 925  
 926  
 927  
 928  
 929  
 930  
 931  
 932  
 933  
 934  
 935  
 936  
 937  
 938  
 939  
 940  
 941  
 942  
 943  
 944  
 945  
 946  
 947  
 948  
 949  
 950  
 951  
 952  
 953  
 954  
 955  
 956  
 957  
 958  
 959  
 960  
 961  
 962  
 963  
 964  
 965  
 966  
 967  
 968  
 969  
 970  
 971  
 972  
 973  
 974  
 975  
 976  
 977  
 978  
 979  
 980  
 981  
 982  
 983  
 984  
 985  
 986  
 987  
 988  
 989  
 990  
 991  
 992  
 993  
 994  
 995  
 996  
 997  
 998  
 999  
 1000

# RESULT 2

GTR5\_HUMAN STANDARD; PRT; 501 AA.  
 ID P22732; Q14770;  
 AC 01-AUG-1991 (Rel. 19, Created)  
 DT 01-AUG-1991 (Rel. 19, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Solute carrier family 2, facilitated glucose transporter, member 5  
 DE (Glucose transporter type 5, small intestine) (Fructose transporter).  
 GN SLC2A5 OR GLUT5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Jejunum;  
 RX MEDLINE=90330675; PubMed=1695905;  
 RA Kayano T., Barant C.F., Fukumoto H., Gould G.W., Fan Y.-S., Eddy R.L.,  
 BYERS M.G., Shows T.B., Seino S., Bell G.I.;  
 RT "Human facilitative glucose transporters. Isolation, functional  
 RT characterization, and gene localization of cDNAs encoding an isoform  
 RT (GLUT5) expressed in small intestine, kidney, muscle, and adipose  
 RT tissue and an unusual glucose transporter pseudogene-like sequence  
 RT (GLUT6).";  
 RL J. Biol. Chem. 265:13276-13282(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Fetal liver;  
 RX Takeda J., Minokoshi Y., Yasuda K., Kayano T., Bell G.I.;  
 RA "Evolution of facilitative sugar transporter gene family:  
 RT characterization of mouse GLUT3 and human GLUT5 genes";  
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymph;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Frange C.,  
 RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length  
 RT human and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [4]  
 RP MASS SPECTROMETRY.  
 RC TISSUE=Breast Cancer;  
 RX MEDLINE=21829512; PubMed=11840567;  
 RA Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Hexath A.,  
 RA Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J.,  
 RA Zvelebil M.J.;  
 RT "Cluster analysis of an extensive human breast cancer cell line  
 RT protein expression map database";  
 RL Proteomics 2:212-223(2002).  
 CC -!- FUNCTION: Cytochalasin B-sensitive carrier. Seems to function  
 CC primarily as a fructose transporter.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- TISSUE SPECIFICITY: Expressed in small intestine, and at much

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:00:52 ; Search time 18 Seconds  
(without alignments)  
1628.639 Million cell updates/sec

Title: US-09-981-947B-2

Perfect score: 2903  
Sequence: 1 MARKQNRNRSKELGLVPLTDD.....NTAMTQAAATTTATKKEHPL 563

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2719	93.7	540	1	GTR9 HUMAN
2	1122.5	38.7	501	1	GTR5 HUMAN
3	1112.5	38.3	502	1	GTR5 RAT
4	1099.5	37.9	501	1	GTR5 MOUSE
5	999	34.4	496	1	GTR11 HUMAN
6	949.5	32.7	492	1	GTR1 BOVIN
7	946.5	32.6	490	1	GTR1 CHICK
8	942.5	32.5	496	1	GTR3 CHICK
9	940.5	32.4	492	1	GTR1 RABIT
10	936.5	32.3	492	1	GTR1 HUMAN
11	929.5	32.0	492	1	GTR1 MOUSE
12	877.5	30.2	451	1	GTR1 RAT
13	877.5	30.2	451	1	GTR1 PIG
14	874	30.1	509	1	GTR4 HUMAN
15	866	29.8	494	1	GTR3 SHEEP
16	864	29.8	494	1	GTR3 BOVIN
17	854	29.4	509	1	GTR4 RAT
18	853	29.4	509	1	GTR4 MOUSE
19	844.5	29.1	493	1	GTR3 MOUSE
20	841	29.0	495	1	GTR3 CANFA
21	841	29.0	509	1	GTR4 BOVIN
22	830	28.6	496	1	GTR3 HUMAN
23	823.5	28.4	486	1	GTR5 RABIT
24	816.5	28.1	493	1	GTR3 RAT
25	772	26.6	522	1	GTR2 RAT
26	767.5	26.4	523	1	GTR2 MOUSE
27	750.5	25.9	533	1	GTR2 CHICK
28	737.5	25.4	390	1	GTR1 SHEEP
29	695.5	24.0	524	1	GTR2 HUMAN
30	679.5	23.4	400	1	GTR3 RABIT
31	480	16.5	488	1	YB91 YEAST
32	457	15.7	204	1	GTR5 BOVIN
33	438	15.1	461	1	CSBC_BACSU

34 413 14.2 464 1 ARAE\_BACSU  
35 411.5 14.2 629 1 MYCT\_HUMAN  
36 410.5 14.1 507 1 GTR6 HUMAN  
37 409.5 14.1 484 1 GALP\_ECOLI  
38 407.5 14.0 547 1 GTR1 LEIDO  
39 400 13.8 468 1 GLCP\_SYNY3  
40 398 13.7 763 1 RGT2 YEAST  
41 396.5 13.7 477 1 GTR8 MOUSE  
42 394.5 13.6 477 1 GTR8 HUMAN  
43 389.5 13.4 510 1 HEXE\_RICCO  
44 389 13.4 540 1 HXTD\_YEAST  
45 389 13.4 592 1 HXT5\_YEAST

#### ALIGNMENTS

##### RESULT 1

GTR9 HUMAN  
ID GTR9 HUMAN STANDARD; PRT; 540 AA.  
AC Q9NRMO;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Solute carrier family 2, facilitated glucose transporter, member 9  
DE (Glucose transporter type 9).  
GN SLC2A9 OR GLUT9.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND VARIANTS GLY-25, ILE-282 AND PRO-350.  
RC TISSUE=Kidney;  
RX MEDLINE=20318624; PubMed=10860667;  
RA Phay J.E., Hussain H.B., Moley J.F.;  
RT "Cloning and expression analysis of a novel member of the facilitative glucose transporter family, SLC2A9 (GLUT9).";  
RL Genomics 66:217-220(2000).  
CC -!- FUNCTION: Facilitative glucose transporter (By similarity).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -!- TISSUE SPECIFICITY: Highly expressed in kidney followed by liver; also detected in placenta, lung, blood leukocytes, heart and skeletal muscle.  
CC -!- SIMILARITY: BELONGS TO THE SUGAR TRANSPORTER FAMILY. GLUCOSE TRANSPORTERS SUBFAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation. The European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL; AF210317; AAF85942.1; -  
CC Genew; HGNC:13446; SLC2A9.  
CC MIM; 606142; -  
CC GO; GO:0016021; C:integral to membrane; NAS.  
CC GO; GO:0005351; P:sugar porter activity; NAS.  
CC GO; GO:0015758; P:glucose transport; NAS.  
CC InterPro; IPR007114; MFS.  
CC InterPro; IPR005828; Sub\_transporter.  
CC InterPro; IPR005829; Sug\_transporter.  
CC InterPro; IPR003663; Sugar\_transpt.  
CC Pfam; PF00083; sugar tr; 1.  
CC PRINTS; PR00171; SUGTRNSPORT.  
CC TIGRfams; TIGR00879; SP; 1.  
CC PROSITE; PS00850; MFS; 1.  
CC PROSITE; PS00216; SUGAR\_TRANSPORT\_1; 1.  
CC PROSITE; PS00217; SUGAR\_TRANSPORT\_2; 1.  
KW Transport; Sugar transport; Transmembrane; Glycoprotein;  
KW Multigene family; Polymorphism.



Db 438 QYVADRMGPVYVLLFAVLLGFFFTFTFLKVPETGRKTFDQISAAPRTPSLLEQEVK 494

Job time : 22 secs

## RESULT 15

S53322  
fructose transport protein GLUT5, intestinal - rabbit  
N;Alternate names: facilitative glucose transporter GLUT5; fructose transporter GLUT5  
C;Species: Oryctolagus cuniculus (domestic rabbit)  
C;Date: 28-Oct-1996 #sequence\_revision 07-Feb-1997 #text\_change 20-Jun-2000  
C;Accession: S53322; S71039  
R;Miyamoto, K.; Tatsumi, S.; Morimoto, A.; Minami, H.; Yamamoto, H.; Sone, K.; Taketani, Biochem. J. 303, 877-883, 1994  
A;Title: Characterization of the rabbit intestinal fructose transporter (GLUT5).  
A;Reference number: S53322; MUID:95071304; PMID:7980458  
A;Accession: S53322  
A;Status: nucleic acid sequence not shown  
A;Molecule type: mRNA  
A;Residues: 1-487 <MIY>  
A;Cross-references: GB:D26482  
R;Miyamoto, K.I.  
submitted to the EMBL Data Library, January 1994  
A;Reference number: S71039  
A;Accession: S71039  
A;Molecule type: mRNA  
A;Residues: 1-424,426-487 <MIW>  
A;Cross-references: EMBL:D26482; NID:G624911; PIDN:BAA05492.1; PID:G633998  
C;Genetics:  
A;Gene: GLUT5  
C;Superfamily: glucose transport protein  
C;Keywords: glycoprotein; sugar transport; transmembrane protein

Query Match 28.9%; Score 838; DB 2; Length 487;  
Best Local Similarity 39.0%; Pred. No. 7.2e-59;  
Matches 193; Conservative 91; Mismatches 181; Indels 30; Gaps 13;  
  
QY 45 GRRRKDWCSLLVA--SLAGAFSGSFLYGYNLVVNAPTPIYKAFYNESWERRHGRPID 102  
DB 5 GQEKKEGRITLVALRTLIAAFGSSFOYAVNVVNCSPSELMTBFYNDTYDRTGELID 64  
QY 103 PDTLTLLWSVTGIFPAIGLGVTLIVKMGKVLGRKHTLLANNNGPAISALLMACSLOAG 162  
DB 65 EPELTLLWSVTVMFSPSGGFGAGLLVGPLVNFGRKGLLFNNFISIVPAIIMGCKVKAR 124  
QY 163 APFMLIVGPIGMDIGGVALSIVLPMYLSISPKETRGSLGVQVTAIFIQIGVFTGQLGLP 222  
DB 125 SFELIILSRLLVIGICAGVSNVVPYVGLAPNIRGALGVESQLFTILGILVAQIFGL- 183  
QY 223 ELGKESWPFYFGVIVVPAVQVLSLPFLPDSRYLLLEKNEARAYKAFQTFLGKADY 282  
DB 184 RSIRQKQWPIILGLTGPAAA--ACPFPPSPRYLLIGQPRCRQ-KALQSLRGWDSV 240  
QY 283 SQVEEVLAESHVORIRLVSVLELLRAPYVWQVVTVVTMACYQLGGLNAIWFYINSI 342  
DB 241 DRELEIRREDEAARAAGLVSVRALCAMEGLAWQLISVVPLM-WQQLSGVNAIYY-DQI 298  
QY 343 FGKAGIPP--AKIPYVTLGTGTETLAAVPSGLVIEHLRRRPLLIGPGMLGPFGLTI 400  
DB 299 Y-----LSPDLDTQYTAATAGVNVLMVCTVVFVESWARLLLLL-GFSPAPTCCULTA 353  
QY 401 TLTLODHAPVYVLSIVGILAIASFCSPGPGIPFILTGEFFQSQRPAAFIAGTVNWL 460  
DB 354 ALALQDVSMPYISVICIIVVIGHAIGP-AIRSLYT-EIFLQSGRPPTW--MGQVHVL 409  
QY 461 SNFAVGLLFPFIQKSLDTYCLVFAFICITGAIYLYFVLPETKNTYAEISQAFSKNKA 520  
DB 410 SNFTVGLVFPFLQWALGLYSFIIFGVACLSITTVYFLIVPETKGSFIEIIRFIRNKV 469  
QY 521 -----YYPE 524  
DB 470 EVSPDREELKDFPPD 484

Search completed: April 6, 2004, 10:18:58

Query Match 29.4%; Score 854; DB 2; Length 509;  
Best Local Similarity 37.9%; Pred. No. 4.1e-60;  
Matches 181; Conservative 99; Mismatches 185; Indels 12; Gaps 6;

QY 57 LVASLAGAFSSFLGYNLVSVNAPTPYIKAFYNESWERRHGRPIDDP-----TTLTWS 111  
DB 24 LVAVFSAVLGSLQFGYNGVINAQKVIEQSYNATWLGROG-PGGPDSIPQGTTLTWA 82

QY 112 VTVSFAIGLVGTIVKMGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGR 171  
DB 83 LSVAFSVGGMISFLIGIISQWLGRKRAMLANNVLAIVGLGALMLANAAASYEILILGR 142

QY 172 FIMGIDGVALSVLPMYLSLSEIPKIRSGISGOVTAIFIGVFTGQLLPELIGKSTW 231  
DB 143 FLIGAYSGLTSGLVPMYVGEIAPTHLRGALGTNLQIAIVIGILVAQVLGLESMLGTATLW 202

QY 232 PYLFGVIVPAVQVQLSLPFLPDSPPRYLLEKHNARAVKAPQTFGLKADVSQVEEVL 291  
DB 203 PLLALTVLPALLQLLPPCPSPRYLYIIRNLEGPARKSLKRLTGWADVSDALAEKD 262

QY 292 ESHVQSRIRLVSVLELLRAPYRVQVTVITMACYQVLCGLNAINFYNSIFGKAGI-PP 350  
DB 263 EKKLERERPLSLQLLGSRTROPLIAAVVLQSLQSGINAVFYSTISFELAGVEQP 322

QY 351 AKIPYVTLSTGGIETLAAVFSLVIEHLGRRL-LIGGFLMGLFFGTTLTITLQDHAP 409  
DB 323 A---YATIGAGVNTVFTLVSVLLVERAGRRTLHLGLAGMCGCAI-LMTVALLLRRVP 378

QY 410 WPYLSIVGILAIASFCSGGIPFLITGEPFQOSORPAAFIAGTVNMLSNFVGLLPPFIQ 469  
DB 379 SMSYVSIVAIFGVAFFGIPGWFIIVAEFVSQGPRAAVAGPSNWTNCFIVGNGF 438

QY 470 PFIQKSLDTCFLVFATCITGAIYLYFVLPEPKRTYAEISQAFSKRNKAYPPEEK 526  
DB 439 QYVADANGPVFLFAVLLGFFITFLVPEPRTGRTFQISATPRTSLLQEYK 495

RESULT 13  
A41751  
Glucose transport protein 3 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 24-Sep-1999  
C:Accession: A41751; S18128  
R:Nagamatsu, S.; Kornhauser, J.M.; Burant, C.F.; Seino, S.; Mayo, K.E.; Bell, G.I.  
J. Biol. Chem. 267, 467-472, 1992  
A:Title: Glucose transporter expression in brain. cDNA sequence of mouse GLUT3, the brain  
A:Reference number: A41751; MUID:92112695; PMID:1730609  
A:Accession: A41751  
A:Molecule type: mRNA  
A:Residues: 1-493 <NAG>  
A:Cross-references: GB:X61093; NID:g51088; PIDN:CAA43406.1; PID:g51089  
C:Genetics:  
A:Gene: GLUT3  
C:Superfamily: glucose transport protein  
C:Keywords: glycoprotein; transmembrane protein  
F:43/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 29.1%; Score 844.5; DB 2; Length 493;  
Best Local Similarity 38.8%; Pred. No. 2.2e-59;  
Matches 179; Conservative 95; Mismatches 182; Indels 5; Gaps 4;

QY 55 SLIVASLAGAFSSFLGYNLVSVNAPTPYIKAFYNESWERRHGRPIDDPDTLLMSVTV 114  
DB 9 SLVFAVTVATIG-SFQFGYNGTVINAPETILKDFLNYTLBERLEDLPSEGLLTALWSLCV 67

QY 115 SIFAIGLVTLIVKMGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFTM 174  
DB 68 AIFSVMGIMGSFVGLVFNFRGRNMLNLAITAGCLMGFAKTAESVEMILGRLLI 127

QY 175 GIDGGVALSVLPMYLSLSEIPKIRSGISGOVTAIFIGVFTGQLLPELIGKSTW 234  
DB 128 GIFCGLCTGVPMIGVSVTAIRGAFGTNLQIGIVGVILVAQIFGLDFILSSEELWPL 187

QY 235 FGIVVPAVQVQLSLPFLPDSPPRYLLEKHNARAVKAPQTFGLKADVSQVEEVLAE 294  
DB 188 LGTLIPAILQSAALPCPSPRFLINKKEEDQATEILQLRWGTSVVQEIQEMKDES 247

QY 295 VQSRIRLVSVLELLRAP-YRVQVTVITMACYQVLCGLNAINFYNSIFGKAGI 353  
DB 248 RMSQEKQVTVLEFRSNVYQVLLISIVLQLS-QQLSGINAVFYSTIGIFKADG--QEP 304

QY 354 PYVTLSTGGIETLAAVFSLVIEHLGRRL-LIGGFLMGLFFGTTLTITLQDHAPV 413  
DB 305 IYATIGAGVNTVFTLVSVLLVERAGRRTLHLGLGMAVCSVFMVTSLLKDDYEMS 364

QY 414 LSVIGILAIASFCSGGIPFLITGEPFQOSORPAAFIAGTVNMLSNFVGLLPPFIQ 473  
DB 365 VCIIVAILLYVAFTEIGPPIWFIIVAEFVSQGPRAAVAGCONWTSNFTLVGLMPSAA 424

QY 474 KSLDTCFLVFATCITGAIYLYFVLPEPKRTYAEISQAF 514  
DB 425 AVLGAIVFIIFAAFLIFLFTFFKVPETKRTFEDIARAF 465

RESULT 14  
B30310  
Glucose transport protein GP2 - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 08-Dec-1989 #sequence\_revision 08-Dec-1989 #text\_change 04-Sep-1998  
C:Accession: B30310  
R:Kaestner, K.H.; Christy, R.J.; McLenithan, J.C.; Braiterman, L.T.; Cornelius, P.; Pek  
Proc. Natl. Acad. Sci. U.S.A. 86, 3150-3154, 1989  
A:Title: Sequence, tissue distribution, and differential expression of mRNA for a putat  
A:Reference number: A30310; MUID:89240694; PMID:2654938  
A:Accession: B30310  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-508 <KAE>  
C:Superfamily: glucose transport protein  
C:Keywords: transmembrane protein

Query Match 28.9%; Score 839.5; DB 2; Length 508;  
Best Local Similarity 37.7%; Pred. No. 5.8e-59;  
Matches 180; Conservative 99; Mismatches 185; Indels 13; Gaps 7;

QY 57 LVASLAGAFSSFLGYNLVSVNAPTPYIKAFYNESWERRHGRPIDDP-----TTLTWS 111  
DB 24 LVAVFSAVLGSLQFGYNGVINAQKVIEQSYNATWLGROG-PGGPDSIPQGTTLTWA 82

QY 112 VTVSFAIGLVGTIVKMGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGR 171  
DB 83 LSVAFSVGGMISFLIGIISQWLGRKRAMLANNVLAIVGLGALMLANAAASYEILILGR 142

QY 172 FIMGIDGVALSVLPMYLSLSEIPKIRSGISGOVTAIFIGVFTGQLLPELIGKSTW 231  
DB 143 FLIGAYSGLTSGLVPMYVGEIAPTHLRGALGTNLQIAIVIGILVAQVLGLESMLGTATLW 202

QY 232 PYLFGVIVPAVQVQLSLPFLPDSPPRYLLEKHNARAVKAPQTFGLKADVSQVEEVL 291  
DB 203 PLLALTVLPALLQLLPPCPSPRYLYIIRNLEGPARKSLKRLTGWADVSDALAEKD 262

QY 292 ESHVQSRIRLVSVLELLRAPYRVQVTVITMACYQVLCGLNAINFYNSIFGKAGI-PP 350  
DB 263 EKKLERERPLSLQLLGSRTROPLIAAVVLQSLQSGINAVFYSTISFELAGVEQP 322

QY 351 AKIPYVTLSTGGIETLAAVFSLVIEHLGRRL-LIGGFLMGLFFGTTLTITLQDHAP 409  
DB 323 A---YATIGAGVNTVFTLVSVLLVERAGRRTLHLGLAGMCGCAI-LMTVALLLRRVP 378

QY 410 WPYLSIVGILAIASFCSGGIPFLITGEPFQOSORPAAFIAGTVNMLSNFVGLL 469  
DB 379 SMSYVSIVAIFGVAFFGIPGWFIIVAEFVSQGPRAAVAGPSNWTNCFIVGNGF 437

QY 470 PFIQKSLDTCFLVFATCITGAIYLYFVLPEPKRTYAEISQAFSKRNKAYPPEEK 526  
DB 439 QYVADANGPVFLFAVLLGFFITFLVPEPRTGRTFQISATPRTSLLQEYK 495

QY 387 GFGLMGLFFGTTITLTDHAPWPPYLSIVGILAIISFSGPGGIPFILTGTBFFQSQ 446  
DB 300 LAGMACCAV-LMTIALALLEQLPMMYSILVAIFGVFAFFVFGPPIFWFVIAELFSQGP 358  
QY 447 RPAATIAATVWNLGNFAVGLLFPFIQKSLDPCFLVPATICITGAIYLYFVLPTKNT 506  
DB 359 RPAATIAVAGFNSWTSNFIWGMCFQVBEQCGPYFIITVLLVFFIFTYFKVPETKRT 418  
QY 507 YABISOAF-----SKENKAYPPEE 525  
DB 419 FDEIASGFQGGASQSDKT--PEE 440  
RESULT 11  
A33801  
muscle-fat glucose-transporter GLUT4 - human  
N;Alternate names: glucose transport protein, insulin-regulated, solute carrier family 2  
C;Species: Homo sapiens (man)  
C;Date: 09-Mar-1990 #sequence\_revision 09-Mar-1990 #text\_change 24-Sep-1999  
R;Accession: A49158; A33801; I55358; I37425  
R;Buse, J.B.; Yasuda, K.; Lay, T.P.; Seo, T.S.; Olson, A.L.; Pessin, J.E.; Karam, J.H.;  
Diabetes 41, 1436-1445, 1992  
A;Title: Human GLUT4/muscle-fat glucose-transporter gene. Characterization and genetic v  
A;Reference number: A49158; MUID:93012518; PMID:1397719  
A;Accession: A49158  
A;Molecule type: DNA  
A;Residues: 1-509 <BUS>  
A;Experimental source: fetal liver  
A;Note: sequence extracted from NCBI backbone (NCBIN:117256, NCBIPI:117257)  
R;Fukunoto, H.; Kayano, T.; Buse, J.B.; Edwards, Y.; Pilch, P.F.; Bell, G.I.; Seino, S.  
J. Biol. Chem. 264, 7776-7779, 1989  
A;Title: Cloning and characterization of the major insulin-responsive glucose transporter  
A;Reference number: A33801; MUID:89255193; PMID:2656669  
A;Accession: A33801  
A;Molecule type: mRNA  
A;Residues: 1-509 <FUK>  
A;Cross-references: GB:M20747; NID:9186552; PIDN:AAAS9189.1; PID:9307076  
R;Buse, J.B.; Yasuda, K.; Lay, T.P.; Seo, T.S.; Liu, M.L.; Olson, A.L.; Pessin, J.E.; Mc  
J. Biol. Chem. 267, 11673-11676, 1992  
A;Title: Expression and regulation of the human GLUT4/muscle-fat facilitative glucose tr  
A;Reference number: I55358; MUID:92291025; PMID:1601840  
A;Accession: I55358  
A;Status: nucleic acid sequence not shown; translation not shown; translated from GB/EMB  
A;Molecule type: DNA  
A;Residues: 1-151,156-509 <RES>  
A;Cross-references: GB:M91463; NID:9183295; PIDN:AAAS2569.1; PID:g183296  
R;Chiaramonte, R.; Martini, R.; Taramelli, R.; Comi, P.  
Gene 130, 307-309, 1993  
A;Title: Identification of the 5' end of the gene encoding a human insulin-responsive gl  
A;Reference number: I37425; MUID:93366193; PMID:7516714  
A;Accession: I37425  
A;Status: preliminary; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-11 <RE2>  
A;Cross-references: EMBL:X58489; NID:9402583; PIDN:CAA41399.1; PID:g402584  
C;Genetics:  
A;Gene: GDB:SLC2A4; GLUT4  
A;Cross-references: GDB:119997; OMIM:138190  
A;Map position: 17p13-17p13  
C;Superfamily: glucose transport protein  
C;Keywords: transmembrane protein  
Query Match 30.1%; Score 874; DB 2; Length 509;  
Best Local Similarity 38.7%; Pred. No. 1e-61;  
Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;  
QY 57 LVASLAGSGSFLYGNLSVNNAPPYKAFYNESWERRHG----REIDPDTLTWSV 112  
DB 24 LVAVFSAVLGSLQFGNIGVFNAPQVKEQSYNTWLGROGPEGPSIPPGTTLTUAL 83  
QY 113 TVSFPAIGGLVGLTIKVKIGLGRKHTLLANNFPAISAAALIMACSLQAGAFEMLIYGRF 172  
DB 84 SVAIFSOGVWISFLIGLISQWLGKRAMLVNNVLAVLGSLMGLANFAASVEMLIYGRF 143

QY 173 IMGIDGGVALSVLPMYLSSEIPKIRBSLQGVTAIFICIGVFTQQLGLQLPILCKESTWP 232  
DB 144 LIGAYSGLTSLVPMYVGEIAPTHLRGALGTNLQIAIVIGLIAQVIGLSLGLTASLWP 203  
QY 233 YLFGVIVPVPVAVVQLSLPPLDSPRYLLEKHENARAVKAFOTELGKADVSQVEEVLAE 292  
DB 204 LLGUTVLPALLQLVLPFCESPRYLIIQNLGSPAKSKRLTGTWADVSGVLAELKDE 263  
QY 293 SHVQSRILSVLELLRAPYVRWQVVIVTMACYQLCGLNAINWFTYNSIFGKAGI--PPA 351  
DB 264 KRKLREERPLSLQLLGSRTTHQPLIIAVVLQSLSGINAVFYYSIFETAGVQCPA 323  
QY 352 KIPYVTLSTGGIETLAAVFSGLVIEHLGRRPL-LIGFGLMGLFFGTLTTLTIQDHAPW 410  
DB 324 ---YATIGAGVVNTVFTLVSVLLVERAGRTTHLGLAGMCGCAI-LMTVALLLERVPA 379  
QY 411 VPLYSIVGLAIIASFCSPGPGIIFLTGTEFFQSQRPAAFIAGTWNLSNFAVGLLFP 470  
DB 380 MGYVSIIVAFGEVAFETGPGPIFWFVIAELFSGPAPAMAVAGPSNWTNFIIGMGFQ 439  
QY 471 FIQKSLDPCFLVPATICITGAIYLYFVLPTKNTYAEISQASKNKAYPPEEK 526  
DB 440 YVAEAMGPVYVLPFALLGLGFFITFLRVBPETGRCTFDQISAAHRTPSLLEQEVK 495  
RESULT 12  
A32101  
glucose transport protein, muscle - rat  
N;Alternate names: insulin-responsive glucose transporter  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 17-Aug-1989 #sequence\_revision 31-Dec-1993 #text\_change 24-Sep-1999  
C;Accession: A32101; S03349; A32387; I55419  
R;Charon, M.J.; Brosius III, F.C.; Alper, S.L.; Lodish, H.F.  
Proc. Natl. Acad. Sci. U.S.A. 86, 2535-2539, 1989  
A;Title: A glucose transport protein expressed predominantly in insulin-responsive tis  
A;Reference number: A32101; MUID:8920363; PMID:2649883  
A;Accession: A32101  
A;Molecule type: mRNA  
A;Residues: 1-509 <CHA>  
A;Cross-references: GB:L04524  
R;James, D.E.; Strube, M.; Mueckler, M.  
Nature 338, 83-87, 1989  
A;Title: Molecular cloning and characterization of an insulin-regulatable glucose tran  
A;Reference number: S03349; MUID:89143771; PMID:2645527  
A;Accession: S03349  
A;Molecule type: mRNA  
A;Residues: 1-509 <JAM>  
A;Cross-references: EMBL:X14771; NID:956501; PIDN:CAA32879.1; PID:g56502  
R;Birbaum, M.J.  
Cell 57, 305-315, 1989  
A;Title: Identification of a novel gene encoding an insulin-responsive glucose transpo  
A;Reference number: A32387; MUID:89195241; PMID:2649253  
A;Accession: A32387  
A;Molecule type: mRNA  
A;Residues: 1-348, Q, 350-509 <BIR>  
A;Cross-references: GB:M25482; NID:9537942; PIDN:AAA41451.1; PID:g537943  
R;Liu, M.L.; Olson, A.L.; Edgington, N.P.; Moye-Rowley, W.S.; Pessin, J.E.  
J. Biol. Chem. 269, 28514-28521, 1994  
A;Title: Myocyte enhancer factor 2 (MEF2) binding site is essential for C2C12 myotube--  
A;Reference number: I55419; MUID:95050643; PMID:7545962  
A;Accession: I55419  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-50 <RES>  
A;Cross-references: GB:L36125; NID:g536863; PIDN:AAA65751.1; PID:g536864  
C;Genetics:  
A;Gene: GLUT4  
A;Introns: 11/3  
C;Superfamily: glucose transport protein  
C;Keywords: transmembrane protein

Query Match 32.3%, Score 930.57, Pred. No. 1e-66;  
Best Local Similarity 39.6%; Conservative 98; Mismatches 178; Indels 11; Gaps 5;  
Matches 188;

QY 57 LVASLAGAGSSFLYGYNLVNVNAPTYIKAFYNESWERHGRPIDPDTLTLLMSVTVSI 116  
DB 12 LMLAVGAGVLSGLOFGYNTGVINAPQKVEIEFYNTQTVHRYGSIILFTLTLLMSLSVAI 71  
QY 117 FAIGLVGTTLVVMIGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRIMGI 176  
DB 72 FSVGGMGSPVGLFVNRFGFRNSMLMMNLLAFVSAVLMGFSKLGKSFEMLIILGRFIIGV 131  
QY 177 DGGVALSVLPMYLSSEISPKIRSGVTAIFICIGVFTGQLLGLPELIGKSTWPLYFG 236  
DB 132 YCGLTTGFPVPMYGEVSPFALRGALGTQLHQLGIVGILIAQVFGDLSIMGNKDLWPLLS 191  
QY 237 VIVVPAVQVLLSLPFLPDSRYLLEKHNARAVKAFQTLGKADYSQVEEVLAEHVQ 296  
DB 192 IIFPALLQCIILVFPFESPRFLINNEENRAKSVLKKLRGTADVTHDQENKESRQM 251  
QY 297 RSIRLVSLELLRAPYVRWQVTVITMACYQLCGLNAINFYNSIFGKAGIPPAKIPV 356  
DB 252 MREKVVILELFRSPAYRQPIILIAVLQSLQSGINAVFYSTISIFKAGV--QQPVYA 309  
QY 357 TLSTGGIETLAAVPSGLVIEHLGRPL-LIGGFLMGLFPGTITITLTQDHAPWVPLY 415  
DB 310 TIGSGIVNTAFTVVSLEFVVERAGRRTLHLGLAGMAGCAI-LMTIALALLEQLPMMSYLS 368  
QY 416 IVGLTALIAFSCSPGGIPILLTGERFQSQORPAAFIAGTVNVLNSNAVGLLFPFIQKS 475  
DB 369 IVAIFGVFAFVFGPGIPFWIVAEFSGQPRPAATAVAGFSNWTNFIIVGMCQFVQV 428  
QY 476 LDTYCLVFAFICITGAIYLYFVLPEKNTYAEISQAF-----SKRNKAYPPEE 525  
DB 429 CGPVFIIFTVLLVLPFIYFKVPTKGTPTDEIASGFRQGGASQSDKT--PEE 481

RESULT 9  
A25949  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Aug-1987 #sequence\_revision 20-Aug-1987 #text\_change 24-Sep-1999  
C:Accession: A25949; I55274  
R: Birnbaum, M.J.; Haepel, H.C.; Rosen, O.M.  
Proc. Natl. Acad. Sci. U.S.A. 83, 5784-5788, 1986  
A:Title: Cloning and characterization of a cDNA encoding the rat brain glucose-transporter  
A:Reference number: A25949; MUID:86287338; PMID:3016720  
A:Accession: A25949  
A:Molecule type: mRNA  
A:Residues: 1-492 <BIR>  
A:Cross-references: GB:M13979; NID:G204413; PIDN:AAA41248.1; PID:G204414  
A:Experimental source: brain  
P: Williams, S.A.; Birnbaum, M.J.  
J. Biol. Chem. 263, 19513-19518, 1988  
A:Title: The rat facilitated glucose transporter gene.  
A:Reference number: I55274; MUID:89066774; PMID:3198639  
A:Accession: I55274  
A:Status: translated from GB/ENBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-492 <RES>  
A:Cross-references: GB:M2063; NID:G204531; PIDN:AAA41297.1; PID:G204533  
A:Experimental source: brain  
C:Genetics:  
A:Introns: 6/3; 38/3; 92/2; 172/3; 227/1; 289/3; 324/3; 358/3; 426/3  
C:Superfamily: glucose transport protein  
C:Keywords: membrane protein

Query Match 32.0%; Score 929.5; DB 2; Length 492;  
Best Local Similarity 39.4%; Conservative 98; Mismatches 179; Indels 11; Gaps 5;  
Matches 187;

QY 57 LVASLAGAGSSFLYGYNLVNVNAPTYIKAFYNESWERHGRPIDPDTLTLLMSVTVSI 116  
DB 12 LMLAVGAGVLSGLOFGYNTGVINAPQKVEIEFYNTQTVHRYGSIILFTLTLLMSLSVAI 71  
QY 117 FAIGLVGTTLVVMIGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRIMGI 176  
DB 72 FSVGGMGSPVGLFVNRFGFRNSMLMMNLLAFVSAVLMGFSKLGKSFEMLIILGRFIIGV 131  
QY 177 DGGVALSVLPMYLSSEISPKIRSGVTAIFICIGVFTGQLLGLPELIGKSTWPLYFG 236  
DB 132 YCGLTTGFPVPMYGEVSPFALRGALGTQLHQLGIVGILIAQVFGDLSIMGNKDLWPLLS 191  
QY 237 VIVVPAVQVLLSLPFLPDSRYLLEKHNARAVKAFQTLGKADYSQVEEVLAEHVQ 296  
DB 192 IIFPALLQCIILVFPFESPRFLINNEENRAKSVLKKLRGTADVTHDQENKESRQM 251  
QY 297 RSIRLVSLELLRAPYVRWQVTVITMACYQLCGLNAINFYNSIFGKAGIPPAKIPV 356  
DB 252 MREKVVILELFRSPAYRQPIILIAVLQSLQSGINAVFYSTISIFKAGV--QQPVYA 309  
QY 357 TLSTGGIETLAAVPSGLVIEHLGRPL-LIGGFLMGLFPGTITITLTQDHAPWVPLY 415  
DB 310 TIGSGIVNTAFTVVSLEFVVERAGRRTLHLGLAGMAGCAI-LMTIALALLEQLPMMSYLS 368  
QY 416 IVGLTALIAFSCSPGGIPILLTGERFQSQORPAAFIAGTVNVLNSNAVGLLFPFIQKS 475  
DB 369 IVAIFGVFAFVFGPGIPFWIVAEFSGQPRPAATAVAGFSNWTNFIIVGMCQFVQV 428  
QY 476 LDTYCLVFAFICITGAIYLYFVLPEKNTYAEISQAF-----SKRNKAYPPEE 525  
DB 429 CGPVFIIFTVLLVLPFIYFKVPTKGTPTDEIASGFRQGGASQSDKT--PEE 481

Query Match 30.2%; Score 877.5; DB 2; Length 451;  
Best Local Similarity 39.9%; Conservative 92; Mismatches 164; Indels 11; Gaps 5;  
Matches 177;

QY 88 FYNESWERHGRPIDPDTLTLLMSVTVSIFATGGLVGLTIIVKMGKVLGRKHTLLANNNGF 147  
DB 2 FYNQTLHRYGESISPAITLTLMSLVAIFSVGGMISGFSVGLFVNRFGFRNSMLMMNLL 61  
QY 148 AISAAALMACSLQAGAFEMLIIVGRFTMGIDGVALSVLPMYLSSEISPKIRSGIQTVAI 207  
DB 62 AFISAVLMGFSKLGKSFEMLIILGRFIIGVYCGTGTGTFVPMYGEVSPFALRGALGTQLH 121  
QY 208 FICIGVFTGQLLGLPELIGKSTWPLYFGVIVVPAVQVLLSLPFLPDSRYLLEKHNAR 267  
DB 122 GIVVGILIAQVGLDLSIMGNELWPLLSVIPALLQCVLLPFCFESPRFLINNEEN 181  
QY 268 RAVKAFQTLGKADYSQVEEVLAEHVQSRSLVLELLRAPYVRWQVTVITMACY 327  
DB 182 RAKSVLKKLAGTADVTRDLQEMKESRQMRKVKVITLFLFSAAYRQPIILIAVLQSLQ 241  
QY 328 QLCGLNAINFYNSIFGKAGIPPAKIPYVTVLSTGGIETLAAVPSGLVIEHLGRRL-LIG 386  
DB 242 QLSGINAVFYSTISIFKAGV--QQPVYATIGSGIVNTAFTVVSLEFVVERAGRRTLHLG 299

RESULT 10  
S04223  
C:Species: Sus scrofa domestica (domestic pig)  
C:Date: 28-Feb-1990 #sequence\_revision 28-Feb-1990 #text\_change 24-Sep-1999  
C:Accession: S04223  
R: Weiler-Guettler, H.; Zinke, H.; Moeskel, B.; Frey, A.; Gassen, H.G.  
Biol. Chem. Hoppe-Seyler 370, 467-473, 1989  
A:Title: cDNA cloning and sequence analysis of the glucose transporter from porcine bio  
A:Reference number: S04223; MUID:89302689; PMID:2472815  
A:Accession: S04223  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-451 <WEI>  
A:Cross-references: GB:X17058; NID:G1955; PIDN:CRA434904.1; PID:G1956  
C:Superfamily: Glucose transport protein  
C:Keywords: transmembrane protein

Development 116, 555-561, 1992  
A>Title: Differential screening of a PCR-generated mouse embryo cDNA library: glucose th  
A/Reference number: I48366  
A/Accession: A30797  
A/Molecule type: mRNA  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Residues: 152-237 <RE2>  
A/Cross-references: EMBL:X69697; NID:G312592; PIDN:CAA49367.1; PID:G312593  
C/Genetics:  
A/Gene: GLUT1  
C/Superfamily: glucose transport protein  
C/Keywords: transmembrane protein

Query Match 32.4%; Score 941.5; DB 2; Length 492;  
Best Local Similarity 39.8%; Pred. No. 4.2e-67;  
Matches 189; Conservative 97; Mismatches 178; Indels 11; Gaps 5;

QY 57 LVASLAGAGSGFLYGVNLSVNVNATPYIKAFYNESWERRGRIPDPTLLTLLSVTVSI 116  
DB 12 LMLAVGAGVLGSLQFGYNTGVNAPQKVIIEFYNTWNRHYGEPISTTTLLSLGVAI 71  
QY 117 FAIGLGVGLTIKVMIGKVLGRKHTLLANNFAISAALLMACSLQAGAFEMLIIVGRFMGI 176  
DB 72 FSVGGMIGSFGVLGVNFRGRNSMLMNNLLAFVSAVLMGFSKLAKSFEMLLIGRFIIGV 131  
QY 177 DGGVALSVLPYMLSEISPKIRSGLSQVTAIFICIGVFTGQLLGLPELLGKESWYPLFG 236  
DB 132 YCGLTTGTFVPMYVGEVSPPTALRGALGTLHQLGIIVGILIAQVFGLDISMGNDLWPLLS 191  
QY 237 VIVVPAVOLLSPFLPDSRPVLLLEKNEARAVKAFQTFGLKADYDSQVEEVLAEHVQ 296  
DB 192 VIFPALLQCIILLPCPESPRFLINNEENRAKSVLKLRGNADVTRDLQEMKEBGRQ 251  
QY 297 RSIRLVSVLELLRAPYVRVQVTVITMACYQLCGLNIAWFTYNSIFGKAGIPPAKIPYV 356  
DB 252 MREKVTIILELFRSPAYRQPILSAVVLQSQLSGINAVFYSTISFEKAGV--QQPVYA 309  
QY 357 TLSTGGIETLAAPVGLVIEHGRPLLIIGFGLMGLFFGTGLTITLTLQDHAPWYPLSI 416  
DB 310 TTGSGIVNTAFTVVSFLVVERAGRTLHLIGLAGMACAVMLMTIALALDQLPWWSYLS 369  
QY 416 IVGILAIASFGSGPGIPFILTGTFFQSQORPAAFIAGTVNLSNPAVGLLFFFIQKS 475  
DB 369 IVAIFGVAFVFEVGGPIPFVIAELFSQGRPAAVAGFNSNWTNFIIVGCMFQYVEQL 428  
QY 476 LDTYCLVPATICITGAIYLYFVLPETRKRTYVAEISQAF-----SKRNKAYPPPE 525  
DB 429 CGPYVFIITVLLVFFIYFKVPETKGRTFDEIASGFRQGGASQSDXT--PPE 481

RESULT 7  
A30797  
glucose transport protein - rabbit  
C/Species: Oryctolagus cuniculus (domestic rabbit)  
C/Date: 01-Dec-1989 #sequence\_revision 30-Sep-1991 #text\_change 24-Sep-1999  
C/Accession: A30797  
R/Asano, T.; Shibasaki, Y.; Kasuga, M.; Kanazawa, Y.; Takaku, F.; Akanuma, Y.; Oka, Y.  
Biochem. Biophys. Res. Commun. 154, 1204-1211, 1988  
A/Title: Cloning of a rabbit brain glucose transporter cDNA and alteration of glucose tr  
A/Reference number: A30797; MUID:88309104; PMID:3408493  
A/Accession: A30797  
A/Molecule type: mRNA  
A/Residues: 1-492 <ASA>  
A/Cross-references: EMBL:M21747; NID:G165633; PIDN:AAA1444.1; PID:G165634  
C/Superfamily: glucose transport protein  
C/Keywords: transmembrane protein

Query Match 32.4%; Score 940.5; DB 2; Length 492;  
Best Local Similarity 39.7%; Pred. No. 5e-67;  
Matches 188; Conservative 98; Mismatches 179; Indels 9; Gaps 3;

QY 57 LVASLAGAGSGFLYGVNLSVNVNATPYIKAFYNESWERRGRIPDPTLLTLLSVTVSI 116

DB 12 LMLAVGAGVLGSLQFGYNTGVNAPQKVIIEFYNTWNRHYGEPISTTTLLSLGVAI 71  
QY 117 FAIGLGVGLTIKVMIGKVLGRKHTLLANNFAISAALLMACSLQAGAFEMLIIVGRFMGI 176  
DB 72 FSVGGMIGSFGVLGVNFRGRNSMLMNNLLAFVSAVLMGFSKLAKSFEMLLIGRFIIGV 131  
QY 177 DGGVALSVLPYMLSEISPKIRSGLSQVTAIFICIGVFTGQLLGLPELLGKESWYPLFG 236  
DB 132 YCGLTTGTFVPMYVGEVSPPTALRGALGTLHQLGIIVGILIAQVFGLDISMGNDLWPLLS 191  
QY 237 VIVVPAVOLLSPFLPDSRPVLLLEKNEARAVKAFQTFGLKADYDSQVEEVLAEHVQ 296  
DB 192 VIFPALLQCIILLPCPESPRFLINNEENRAKSVLKLRGNADVTRDLQEMKEBGRQ 251  
QY 297 RSIRLVSVLELLRAPYVRVQVTVITMACYQLCGLNIAWFTYNSIFGKAGIPPAKIPYV 356  
DB 252 MREKVTIILELFRSPAYRQPILSAVVLQSQLSGINAVFYSTISFEKAGV--QQPVYA 309  
QY 357 TLSTGGIETLAAPVGLVIEHGRPLLIIGFGLMGLFFGTGLTITLTLQDHAPWYPLSI 416  
DB 310 TTGSGIVNTAFTVVSFLVVERAGRTLHLIGLAGMACAVMLMTIALALDQLPWWSYLS 369  
QY 416 IVGILAIASFGSGPGIPFILTGTFFQSQORPAAFIAGTVNLSNPAVGLLFFFIQKS 475  
DB 369 IVAIFGVAFVFEVGGPIPFVIAELFSQGRPAAVAGFNSNWTNFIIVGCMFQYVEQL 428  
QY 476 LDTYCLVPATICITGAIYLYFVLPETRKRTYVAEISQAF-----SKRNKAYPPPE 525  
DB 429 CGPYVFIITVLLVFFIYFKVPETKGRTFDEIASGFRQGGASQSDXT--PPE 481

RESULT 8  
A27217  
glucose transport protein - human  
N/Alternate names: monosaccharide transport protein  
C/Species: Homo sapiens (man)  
C/Date: 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 24-Sep-1999  
C/Accession: A27217; A45957; S13065  
R/Mueckler, M.; Caruso, C.; Baldwin, S.A.; Panico, M.; Blench, I.; Morris, H.R.; Allar  
Science 229, 941-945, 1985  
A/Title: Sequence and structure of a human glucose transporter.  
A/Reference number: A27217; MUID:85272595; PMID:3839598  
A/Accession: A27217  
A/Molecule type: mRNA  
A/Residues: 1-492 <WUE>  
A/Cross-references: GB:K03195; NID:G183302; PIDN:AAA52571.1; PID:G183303  
A/Note: Parts of this sequence were confirmed by peptide sequencing  
A/Note: This protein was shown not to be processed at amino or carboxyl ends  
R/Fukunoro, H.; Seino, S.; Imura, H.; Seino, Y.; Bell, G.I.  
Diabetes 37, 657-661, 1988  
A/Title: Characterization and expression of human HepG2/erythrocyte glucose-transporte  
A/Reference number: A45957; MUID:88196648; PMID:2834252  
A/Accession: A45957  
A/Molecule type: DNA  
A/Residues: 1-6 <FUK>  
A/Cross-references: GB:M20653; NID:G183974; PIDN:AA861084.1; PID:G463112  
R/Wadzinski, B.E.; Shanahan, M.F.; Seamon, K.B.; Ruoho, A.E.  
Biochem. J. 272, 151-158, 1990  
A/Title: Localization of the forskolin photolabelling site within the monosaccharide t  
A/Reference number: S13065; MUID:91090708; PMID:2264820  
A/Accession: S13065  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 254-268 <WAD>  
C/Genetics:  
A/Gene: GDB:SLC2A1; GLUT; GLUT1  
A/Cross-references: GDB:120627; OMIM:138140  
A/Map position: 1p35-p31.3  
C/Superfamily: glucose transport protein  
C/Keywords: glycoprotein; transmembrane protein  
F/45/Binding site: carbohydrate (Asn) (covalent) #status experimental  
F/411/Binding site: carbohydrate (Asn) (covalent) #status absent

I45902  
glucose transporter type I - bovine  
C/Species: Bos primigenius taurus (cattle)  
C/Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 24-Sep-1999  
C/Accession: I45902  
R/Boado, R.J.; Pardridge, W.M.  
Mol. Cell. Neurosci. 11, 224-232, 1991  
A/Title: Molecular cloning of the bovine blood-brain barrier glucose transporter cDNA and its complementary DNA  
A/Reference number: I45902  
A/Accession: I45902  
A/Molecule type: mRNA  
A/Status: preliminary; translated from GB/ENBL/DBJ  
A/Residues: 1-492 <BOA>  
A/Cross-references: GB:M60448; NID:G163104; PIDN:AAA30550.1; PID:G163105  
C/Genetics:  
A/Gene: GLUT-1  
C/Superfamily: glucose transport protein

Query Match 32.7%; Score 949.5; DB 2; Length 492;  
Best Local Similarity 40.0%; Pred. No. 9.7e-68;  
Matches 190; Conservative 100; Mismatches 174; Indels 11; Gaps 5;

Qy 57 LVASLAGAFSSFLYGYNLVSVNAPTPIKAFYNESWERRHGRPDPTLTLLWSVTYSI 116  
Db 12 LMLAVGAGVLSLQFGVNTGVINAPQKVIEEFVNTQVRYGEPPTATLTLLSLVAI 71  
Qy 117 FATGGLVGLTVIKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLVGRFNGI 176  
Db 72 FSVGGMIGSPSVGLFVNRFGRRNSMLMNLAFVSAVLMGFSLKGSPEMLILGRFIIIV 131  
Qy 177 DGGVALSVLPMYLSSEISPEIRSLGQVTAIFICIGVFTGQLLGLPELLGKSTWYPLFG 236  
Db 132 YCLGTTGFVPMYGVESPELGRALGTLHOLGIVGVGILIAQVGLDSINGQELMPLLS 191  
Qy 237 VIIVPAVQVLLSPFLPDSPRYLLEKHEARAVKAFQFFLGKADVSQVEVEVLAESHVQ 296  
Db 192 VIFIPALLQCLLPFCPSPRFLINNEENRAKSVLKLRTADVTRDLQEMKEESQRM 251  
Qy 297 RSRLVSVLELLAPVVRVQVTVITVMACYLQGLNNAIWFYTNIFPGKAGIPPAPYV 356  
Db 252 MREKKVTILELPSRAVROPILIAVVLQSLSGINAVFYSTSIPEKAGV--QQPYA 309  
Qy 357 TLSTGGIETLAAVFGSLVIEHLGRPL-LIGFGFLMGLFPFGTLTTLTLQDHAPWPYLS 415  
Db 310 TIGSGIVNTAFTVSVLFVVERAGRTLHLIGLAGMAGCAV-LMTALLALLERLPMWSYLS 368  
Qy 416 IVGILAIASFCSGPGGIPFILTGEFFQOSQORPAAFIAGTVNLSNFAVGLLRPFIOKS 475  
Db 369 IVAIFGVFAFFVGVGPDPFVFAELFSGQPRPAIAVAGFSNWTNFIVGMCFQYVEQL 428  
Qy 476 LDYCYFLVFTATCITGAIYLYFVLPETKRTYAEISQAF-----SKRNKAYPPEE 525  
Db 429 CGSYVFIITVLLVLFIFTYFKVPETKRTFDEIASGPRGGASQSDKT--PEE 481

RESULT 5  
A41264  
glucose transport protein 3 - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 28-May-1992 #sequence\_revision 28-May-1992 #text\_change 04-Sep-1998  
C/Accession: A41264  
R/White, M.K.; Rall, T.B.; Weber, M.J.  
Mol. Cell. Biol. 11, 4448-4454, 1991  
A/Title: Differential regulation of glucose transporter isoforms by the src oncogene in chicken erythrocytes  
A/Reference number: A41264; MUID:91342646; PMID:1875932  
A/Accession: A41264  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-496 <WHI>  
A/Cross-references: GB:M37785  
C/Superfamily: glucose transport protein  
C/Keywords: transmembrane protein

Query Match 32.5%; Score 942.5; DB 2; Length 496;  
Best Local Similarity 39.3%; Pred. No. 3.5e-67;  
Matches 193; Conservative 94; Mismatches 187; Indels 17; Gaps 5;

Qy 49 RKMDSLLVASLAGAFSSFLYGYNLVSVNAPTPIKAFYNESWERRHGRPDPTLTLL 108  
Db 4 KKKITASLIYAVSVAAG-SLQFGVNTGVINAPEKIIQAFYRTUSQSRGETISPELLTS 62  
Qy 109 LWSVTYSIFAIGSLVGLTVIKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLI 168  
Db 63 LWSLSVAIFSVGGMIGSPSVGLFVNRFGRRNSMLLVNLAFFAGGALMALSKIKAEMLI 122  
Qy 169 VGFPMIGDGGVALSVLPMYLSSEISPEIRSLGQVTAIFICIGVFTGQLLGLPELLGKE 228  
Db 123 IGRFIIIGLFCGLCTGFVPMYIIEVSPISLRGAFGLTNOLGIVGVILVAQIFGLEGIMGTE 182  
Qy 229 STWPYLFVIVPAVQVLLSPFLPDSPRYLLEKHEARAVKAFQFFLGKADVSQVEVEE 288  
Db 183 ALWPLLGLFTIIPAVLQCVALLFCPSPRFLINKEBEKAGTVLQKLRGTQDVSQDISE 242  
Qy 289 VLAESHVQSRISRLSVLELLRAPYVRVQVTVITVMACYLQGLNNAIWFYTNISFGKAGI 348  
Db 243 MKESAKMSQEKATVLELFRSPVNRQPIIISITLQSLQSLSGINAVFYSTGPIPERAGI 302  
Qy 349 PPAKIPYVTLSTGGIETLAAVFGSLVIEHLGRPL-LIGFGFLMGLFPFGTLTTLTLQDHA 408  
Db 303 --TQPYATIGAGVNTVFTVSVLSFLVERAGRTLHLVGLGMAVCAAAVNTIALAKE-- 358  
Qy 409 PWPYLSIVGILAIASFCSGPGGIPFILTGEFFQOSQORPAAFIAGTVNLSNFAVGLL 468  
Db 359 KMRYSIVATGCVFALFEIGPDPFVFAELFSGQPRPAIAVAGFSNWTNFIVGM 418  
Qy 469 FPIQSLDYCYFLVFTATCITGAIYLYFVLPETKRTYAEISQAFSKRNKAYPPEEKID 528  
Db 419 FPAEKLGGYVFLFVFLVFLFIFTYFKVPETKRTFEDISRGF-----EEQVE 469  
Qy 529 SAVTDAPASPP 539  
Db 470 ---TSPSPSP 477

RESULT 6  
S09705  
glucose transport protein GT1 - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 12-Feb-1993 #sequence\_revision 12-Feb-1993 #text\_change 17-Nov-2000  
C/Accession: S09705; A30310; I49697; I48366  
R/Reed, B.C.; Shade, D.; Alperovich, F.; Vang, M.  
Arch. Biochem. Biophys. 279, 261-274, 1990  
A/Title: 3T3-L1 adipocyte glucose transporter (HepG2 class): sequence and regulation of expression in 3T3-L1 adipocytes  
A/Reference number: S09705; MUID:90274408; PMID:2190533  
A/Accession: S09705  
A/Molecule type: mRNA  
A/Residues: 1-492 <REE>  
A/Cross-references: EMBL:M22998; NID:G193551; PIDN:AAA37707.1; PID:G309262  
R/Kaestner, K.H.; Christy, R.J.; McLenithan, J.C.; Braiterman, L.T.; Cornelius, P.; Pek  
Proc. Natl. Acad. Sci. U.S.A. 86, 3150-3154, 1989  
A/Title: Sequence, tissue distribution, and differential expression of mRNA for a putative glucose transporter in rat liver and adipose tissue  
A/Reference number: A30310; MUID:89240694; PMID:2654938  
A/Accession: A30310  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-51, 153-192, 194, 196-402, 404-464, 466-492 <KAE>  
R/Murakami, T.; Nishiyama, T.; Shirotani, T.; Shinohara, Y.; Kan, M.; Ishii, K.; Kanai, Y.  
J. Biol. Chem. 267, 9300-9306, 1992  
A/Title: Identification of two enhancer elements in the gene encoding the type 1 glucose transporter  
A/Reference number: I49697; MUID:92250534; PMID:1339457  
A/Accession: I49697  
A/Status: translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-6 <RES>  
A/Cross-references: GB:D10229; NID:G220414; PIDN:BAA01061.1; PID:G3582350  
R/Smith, D.E.; Gridley, T.

Db 197 GLTGVPAALQLLLPFPSPRYLLIQKDEAAAKKALQTLRGWSDVREVAIRQDEA 256  
 QY 296 QRSIRLVSVLELLRAPPYVWQVTVITVMACVQLCGLNAINWYTNISFGKAGIPAKIPY 355  
 Db 257 EKAAGFISVVKLFMRSLRWQLLSITVLVGGQOLSGVNAIYYADQIYLSAGVPEEHVQY 316  
 QY 356 VTLSTGGIB--TLAAVFGSLVTEHLGRRLPILIGFGLMGLFFGTLTTLTLQDHAPWPV 412  
 Db 317 VTAGTGANVWTFPCAVP---VVELLGRRLLLLLGFSCLCIACCVTALALQDTVSWMP 373  
 QY 413 YLSIVGILAIASFCSPGGIPFILTGEFFQOSQORPAAPIIAGTVNWLNSFVAGLLPFI 472  
 Db 374 YISIVCVISVIGHALGPSIPALLITEIFLQSSRPSAFVWGSVHMLSNTVGLRPFI 433  
 QY 473 QKSLDTCFLVPATICITGAIYLVFLPETKRTVABISOAFSKKNK---AYPPEEKI 527  
 Db 434 QEGLPYSFIVFAVICLLTIIYFIVLPETKRTFIEINQIFTKNKVSEVYPEKEEL 491

RESULT 2  
 I53268  
 glut 5 protein - rat  
 N;Alternate names: fructose transporter  
 C;Species: Rattus norvegicus (Norway rat)  
 C;Date: 02-Aug-1996 #sequence\_revision 02-Aug-1996 #text\_change 24-Sep-1999  
 A;Accession: I53268; I51895  
 R;Inukai, K.; Asano, T.; Katagiri, H.; Ishihara, H.; Anai, M.; Fukushima, Y.; Tsukuda, K.  
 Endocrinology 133, 2009-2014, 1993  
 A;Title: Cloning and increased expression with fructose feeding of rat jejunal GLUTs.  
 A;Reference number: I53268; PMID:94008761; PMID:8404647  
 A;Accession: I53268  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-502 <RES>  
 A;Cross-references: GB:D13871; NID:9436578; PIDN:BA002983.1; PID:9436579  
 R;Rand, E.B.; Depauli, A.M.; Davidson, N.O.; Bell, G.I.; Burant, C.F.  
 Am. J. Physiol. 264, G1169-G1176, 1993  
 A;Title: Sequence, tissue distribution, and functional characterization of the rat fructose transporter.  
 A;Reference number: I51895; PMID:93325725; PMID:8333543  
 A;Accession: I51895  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: mRNA  
 A;Residues: 1-163, 'A', 165-284, 'A', 286-451, 'T', 453-478, 'N', 480-502 <RE2>  
 A;Cross-references: GB:I05195; NID:9204415; PIDN:AAA02627.1; PID:9204416  
 C;Genetics:  
 A;Gene: glut 5  
 C;Superfamily: glucose transport protein

Query Match 38.2%; Score 1109.5; DB 2; Length 502;  
 Best Local Similarity 44.7%; Pred. No. 1.8e-80;  
 Matches 217; Conservative 97; Mismatches 167; Indels 5; Gaps 2;

QY 56 LLVASIAGAFSGSFLYGNLSVNVNAPTPIYKAFYNESWERRHGRPIDPDTLLWSVTVS 115  
 Db 16 LALATPLAAFGSSFOYGVNVAVNSPSEFMQOYNDTYDRNKENIESFTLLWSLTVS 75  
 QY 116 ITAIGLVGTLIVKMGKVLGRKHTLLANNPFAISALLMACSLQACAFEMLIYGRFTMG 175  
 Db 76 MPFFGFGISLMVGFVNLNLRGKALLFNFIIPAILMGCSTAKSFETIIASRLLVG 135  
 QY 176 IDGGVALSVLPMLSEISPKIRSGISLQGVTAIFICIGVFTGQLLGLPELLGKSTWPLYF 235  
 Db 136 ICAGISSNVPMYLGELAPKNLRGALGVVPOLFITVGLVLAQLGLRSVLASEGWPILL 195  
 QY 236 GVIVPAAVQLLSLPLDPSPRYLLEKHNEARAVKAFQTLGKADVSQVEEVLAEASHV 295  
 Db 196 GLTGPVAGLQLLLPFPSPRYLLIQKDEAAAKKALQTLRGWSDVREVAIRQDEA 255  
 QY 296 QRSIRLVSVLELLRAPPYVWQVTVITVMACVQLCGLNAINWYTNISFGKAGIPAKIPY 355  
 Db 256 EKAAGFISVVKLFMRSLRWQLLSITVLVGGQOLSGVNAIYYADQIYLSAGVPEEHVQY 315  
 QY 356 VTLSTGGIB--TLAAVFGSLVTEHLGRRLPILIGFGLMGLFFGTLTTLTLQDHAPWPV 415

Db 316 VTAGTGANVWTFWTVFVVELWGRNLLIGFSTCLTACIVLTVLALQNTISWMBYVS 375  
 QY 416 IVGILAIASFCSPGGIPFILTGEFFQOSQORPAAPIIAGTVNWLNSFVAGLLPFI 475  
 Db 376 IVCVIVVIGHAVGPSIPALPITEIFLQSSRPSAYMIGSVHMLSNTVGLIPFIQVQ 435  
 QY 476 LDTYCFVLPATICITGAIYLVFLPETKRTVABISOAFSKKNK---AYPPEEKIDSAVT 532  
 Db 436 LQYFIFIIAILCLLTIIYFVWVVPETKRTFVINOIFAKKNKVDVYP--EKEEKELN 493  
 QY 533 DAPASS 538  
 Db 494 DUPPAT 499

RESULT 3  
 G02864  
 fructose transporter - human  
 C;Species: Homo sapiens (man)  
 C;Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 24-Sep-1999  
 C;Accession: G02864  
 R;Takada, J.  
 submitted to the EMBL Data Library, January 1994  
 A;Reference number: G07991  
 A;Accession: G02864  
 A;Status: preliminary; translated from GB/EMBL/DBJ  
 A;Molecule type: DNA  
 A;Residues: 1-481 <TAX>  
 A;Cross-references: EMBL:U11843; NID:9516513; PIDN:AAB0641.1; PID:9516515  
 C;Genetics:  
 A;Gene: GLUT5  
 A;Introns: 11/2; 24/3; 78/2; 120/1; 171/1; 213/1; 275/3; 312/3; 346/3; 372/1; 414/3  
 C;Superfamily: glucose transport protein

Query Match 35.6%; Score 1032.5; DB 2; Length 481;  
 Best Local Similarity 44.0%; Pred. No. 2.3e-74;  
 Matches 209; Conservative 94; Mismatches 155; Indels 17; Gaps 5;

QY 60 SIAGAFSGSFLYGNLSVNVNAPTPIYKAFYNESWERRHGRPIDPDTLLWSVTVSIPA 118  
 Db 7 SNKEGWTWL-----LSTPQHMLMQOYFNYTYGRTGEFMDFFLTLWSVTVSMFP 59  
 QY 119 IGLVGTLLIVKMGKVLGRKHTLLANNPFAISALLMACSLQACAFEMLIYGRFTMG 178  
 Db 60 FCGFISLVLVGLVPLVNVKFGKALLFNFIIPAILMGCSTAKSFETIIISRLVGICA 119  
 QY 179 GVALSVLPMLSEISPKIRSGISLQGVTAIFICIGVFTGQLLGLPELLGKSTWPLYF 238  
 Db 120 GVSSNVPMYLGELAPKNLRGALGVVPOLFITVGLVLAQLGLRNLNANVDGHPILLGLT 179  
 QY 239 VVPVAVQLLSLPLDPSPRYLLEKHNEARAVKAFQTLGKADVSQVEEVLAEASHV 298  
 Db 180 GVGAALQLLLPFPFPSPRYLLIQKDEAAAKKALQTLRGWSDVREVAIRQDEA 239  
 QY 299 IRLVSVLELLRAPPYVWQVTVITVMACVQLCGLNAINWYTNISFGKAGIPAKIPV 358  
 Db 240 AGFISVVKLFMRSLRWQLLSITVLVGGQOLSGVNAIYYADQIYLSAGVPEEHVQV 299  
 QY 359 STGGIE--TLAAVFGSLVTEHLGRRLPILIGFGLMGLFFGTLTTLTLQDHAPWPV 415  
 Db 300 GTGANVWTFCAVP---VVELLGRRLLLLLGFSCLCIACCVTALALQDTVSWMBYIS 356  
 QY 416 IVGILAIASFCSPGGIPFILTGEFFQOSQORPAAPIIAGTVNWLNSFVAGLLPFI 475  
 Db 357 IVCVIVVIGHALGPSIPALLITEIFLQSSRPSAFVWGSVHMLSNTVGLIPFIQVQ 416  
 QY 476 LDTYCFVLPATICITGAIYLVFLPETKRTVABISOAFSKKNK---AYPPEEKI 527  
 Db 417 LQYFIFVFNICLLTIIYFIVLPETKRTFIEINQIFTKNKVSEVYPEKEEL 471

RESULT 4



	Query Match	38.7%;	Score 1122.5;	DB 2;	Length 501;
	Best Local Similarity	46.4%;	Pred. No. 1.6e-81;		
	Matches 22;	Conservative	92;	Mismatches 155;	Indels 9; Gaps 3;
Qy	56	LVASLAGAFSGSFLXYGNLSVNVNAPPYTKAFVNSWERRHGRIPDDPTLTLLWSVTVS	115		
Db	17	LALATLAAFGSSFOYGYNVAAVNSPALLQGYNFYTGRTGEFMDFFUTLLWSVTVS	76		
Qy	116	IFAIGGLVGTLIWVMIGKVLGRKHTLIANNQFAISALLMACSLQAGAFMLIVGREIMG	175		
Db	77	MEFFGGFIGSLLVGPLVNGKGRGALLFNNI FSVIPAILMGCSRATSPELLIISRLLVG	136		
Qy	176	IDGGVALSVPMYLSLSPEIKSGISQGVTAIFICIGVFTGQLGLPELLCKESTWYLP	235		
Db	137	ICAGSVNVVMYLGELAPKRLKALGVVQPLFTVIGILVAQIFGLRNLIANVDGWILL	196		
Qy	236	GVIVPVAVQQLSLPFLPDSPRYLLLEKHNRAAVKAQTFLGKADYSQVEEVLASHV	295		

APPLICANT: Eos Biotechnology, Inc.  
TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and  
TITLE OF INVENTION: Methods of Screening for Modulators of Cancer  
FILE REFERENCE: 018501-012500US  
CURRENT APPLICATION NUMBER: US/10/295,027  
CURRENT FILING DATE: 2002-11-13  
PRIOR APPLICATION NUMBER: US 09/663,733  
PRIOR FILING DATE: 2000-09-15  
PRIOR APPLICATION NUMBER: US 60/350,666  
PRIOR FILING DATE: 2001-11-13  
PRIOR APPLICATION NUMBER: US 60/335,394  
PRIOR FILING DATE: 2001-11-15  
PRIOR APPLICATION NUMBER: US 60/332,464  
PRIOR FILING DATE: 2001-11-21  
PRIOR APPLICATION NUMBER: US 60/334,393  
PRIOR FILING DATE: 2001-11-29  
PRIOR APPLICATION NUMBER: US 60/340,376  
PRIOR FILING DATE: 2001-12-14  
PRIOR APPLICATION NUMBER: US 60/347,211  
PRIOR FILING DATE: 2002-01-08  
PRIOR APPLICATION NUMBER: US 60/347,349  
PRIOR FILING DATE: 2002-01-10  
PRIOR APPLICATION NUMBER: US 60/355,250  
PRIOR FILING DATE: 2002-03-08  
PRIOR APPLICATION NUMBER: US 60/356,714  
PRIOR FILING DATE: 2002-02-13  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1386  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 308  
LENGTH: 501  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-295-027-308

Query Match 38.7%; Score 1122.5; DB 15; Length 501;  
Best Local Similarity 46.4%; Pred. No. 8.3e-96;  
Matches 222; Conservative 92; Mismatches 155; Indels 9; Gaps 3;  
QY 56 LVASLAGAGSSFLYGYNLSVNVNAPTYKAFYNESWERRHGRPIDPDTTLMSVTVS 115  
DB 17 LALATLIAAFGSSFGYGVNVAVNSPALLMQQFYNETYGRTEGFEDFPLTLMSVTVS 76  
QY 116 IFAIGLVGTLLVKGKVLGRKHTLLANNNGFAISALLMACSLOAGAFEMLIIVGRFIMG 175  
DB 77 MFPGGFTGSLVGLVGNKFGKGLLENNIFSIYPAILMGCSRVATSFELIISRLVVG 136  
QY 176 IDGVALSVLPMYLSLSEISKEIRSGVQTATFICIGVFTGQLGLPELLGKESWPLYLF 235  
DB 137 ICAGVSSNVVPMYLGELAPKNRGLGVVPQFTIVGILVAQIFGRNLLANVDGWPIIL 196  
QY 236 GVIVVPVAVVQLLSLPDPSRYLLLEKHEARAVKAFQTFILGKADVSQVEEVLAEHV 295  
DB 197 GLTGVPALQILLPFFESPRLYLIQKDEAAKALQTLRGWSDVDEVAIEQDEEA 256  
QY 296 QRSIRLSVLELLRAPYVWQVTVVITWACYQLCGLNAINFYTNISIFKAGIPPAKIPY 355  
DB 257 ERAAGFISVLKFRMSLRWQLLSIIVLMGGQQLSGVNAIYYADQIYLSAGVPEEHVQY 316  
QY 356 VTLSGGIE---TLAAVFSGLVIEHLGRPLIIGFGLMGLFFGTLITLTLODHAPWVP 412  
DB 317 VTAGTGANVNVMTFCAVF---VVELLGRLLLLGFSICLIACCCLTAALQDQTVSWMP 373  
QY 413 YLSIVGILAIASFCSGPGGIFILTGEPFQSQRPAAFIAGTVNWLNSNPAVGLLFFPI 472  
DB 374 YISIVCVISYVIGHALGPSPIFALLITEIFLOSSRPSAFWCGSVHLSNFTVGLIFPI 433  
QY 473 QKSLDTYCFVLPATICITGAIYLYFVLPETKNTYAEISQAFSKENK---AYPPEEKI 527  
DB 434 QEGGLGPSYFIVFAVICLLTTIIFLIVPETKAKTIEINQIFTKMKVSEVYPEKEEL 491

; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 88  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-186-88

Query Match 65.8%; Score 1911; DB 12; Length 396;  
Best Local Similarity 99.5%; Pred. No. 1.3e-169;  
Matches 376; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 137 RKHTLLANNNGFAISALLMACSLOAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 196  
Db 11 RKHTLLANNNGFAISALLMACSLOAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 70  
Qy 137 IRGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYLFVGVVPAVVQVLLSLPFLDPS 256  
Db 71 IRGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYLFVGVVPAVVQVLLSLPFLDPS 130  
Qy 257 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 316  
Db 131 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 190  
Qy 317 VVTVIVTMACYQLCGNALWFYNTSIFGKAGIPPAKIPVYTLSTGGIETLAAVFSGLVIE 376  
Db 191 VVTVIVTMACYQLCGNALWFYNTSIFGKAGIPPAKIPVYTLSTGGIETLAAVFSGLVIE 250  
Qy 377 HLGRRPLLIGGGLMGLFPGTITITLQDHAPWPVYLSIVGILAIASFCSGGPGGIPFI 436  
Db 251 HLGRRPLLIGGGLMGLFPGTITITLQDHAPWPVYLSIVGILAIASFCSGGPGGIPFI 310  
Qy 437 LTGEFFQOSORPAAFIAGTVNWLNSFVAVGLLFPFIQKSLDTCFLVFPATICITGAIYLY 496  
Db 311 LTGEFFQOSORPAAFIAGTVNWLNSFVAVGLLFPFIQKSLDTCFLVFPATICITGAIYLY 370  
Qy 497 FVLPTKNTYAEISQAF 514  
Db 371 FVLPTKNTYAEISQAF 388

RESULT 14  
US-10-188-186-82  
; Sequence 82, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814

; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 82  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-186-82

Query Match 59.3%; Score 1721; DB 12; Length 362;  
Best Local Similarity 90.5%; Pred. No. 6.4e-152;  
Matches 342; Conservative 1; Mismatches 1; Indels 34; Gaps 1;  
Qy 137 RKHTLLANNNGFAISALLMACSLOAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 196  
Db 11 RKHTLLANNNGFAISALLMACSLOAGAFEMLIIVGRFIMGIDGGVALSVLPYMLSEISPK 70  
Qy 137 IRGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYLFVGVVPAVVQVLLSLPFLDPS 256  
Db 71 IRGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYLFVGVVPAVVQVLLSLPFLDPS 130  
Qy 257 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 316  
Db 131 RYLLLEKNEARAVKAFQTFGLKADVSQVEVEVLAEHVORSIRLVSVLELLRAPYVRW 190  
Qy 317 VVTVIVTMACYQLCGNALWFYNTSIFGKAGIPPAKIPVYTLSTGGIETLAAVFSGLVIE 376  
Db 191 VVTVIVTMACYQLCGNALWFYNTSIFGKAGIPPAKIPVYTLSTGGIETLAAVFSGLVIE 245  
Qy 377 HLGRRPLLIGGGLMGLFPGTITITLQDHAPWPVYLSIVGILAIASFCSGGPGGIPFI 436  
Db 246 -----DHAPWPVYLSIVGILAIASFCSGGPGGIPFI 276  
Qy 437 LTGEFFQOSORPAAFIAGTVNWLNSFVAVGLLFPFIQKSLDTCFLVFPATICITGAIYLY 496  
Db 277 LTGEFFQOSORPAAFIAGTVNWLNSFVAVGLLFPFIQKSLDTCFLVFPATICITGAIYLY 336  
Qy 497 FVLPTKNTYAEISQAF 514  
Db 337 FVLPTKNTYAEISQAF 354

RESULT 15  
US-10-295-027-308  
; Sequence 308, Application US/10295027  
; Publication No. US20030232350A1  
; GENERAL INFORMATION:  
; APPLICANT: Afar, Daniel  
; APPLICANT: Aziz, Natasha  
; APPLICANT: Ginsberg, Wendy M.  
; APPLICANT: Gish, Kurt C.  
; APPLICANT: Glynn, Richard  
; APPLICANT: Hevez, Peter A.  
; APPLICANT: Mack, David H.  
; APPLICANT: Murray, Richard  
; APPLICANT: Watson, Susan R.

; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 86  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-186-86

Query Match 66.4%; Score 1929; DB 12; Length 396;  
Best Local Similarity 100.0%; Pred. No. 2.7e-171; Indels 0; Gaps 0;  
Matches 378; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 137 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 196  
DB 11 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 70  
QY 197 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLFGVIVPAVQVLLSLPFLPDS 256  
DB 71 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLFGVIVPAVQVLLSLPFLPDS 130  
QY 257 RYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRW 316  
DB 131 RYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRW 190  
QY 317 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLV 376  
DB 191 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLV 250  
QY 377 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 436  
DB 251 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 310  
QY 437 LTGEFFQOSORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFATICTGAIYLY 496  
DB 311 LTGEFFQOSORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFATICTGAIYLY 370  
QY 497 FVLPEKNTYAEISQAF 514  
DB 371 FVLPEKNTYAEISQAF 388

RESULT 12  
US-10-188-186-84  
; Sequence 84, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046

; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 84  
; LENGTH: 396  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-186-84

Query Match 66.1%; Score 1920; DB 12; Length 396;  
Best Local Similarity 99.5%; Pred. No. 1.9e-170; Indels 0; Gaps 0;  
Matches 376; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
QY 137 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 196  
DB 11 RHTLLANNFAISALLMACSLQAGAFEMLIYGRFIMGIDGGVALSVLPMYLSEISPK 70  
QY 197 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLFGVIVPAVQVLLSLPFLPDS 256  
DB 71 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWPLFGVIVPAVQVLLSLPFLPDS 130  
QY 257 RYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRW 316  
DB 131 RYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEHVQSRIRLSVLELLRAPYVRW 190  
QY 317 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLV 376  
DB 191 VVTIVTMACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLSTGGIETLAAVFSGLV 250  
QY 377 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 436  
DB 251 HLGRRPLLIGGFGMLGFFGTLTITLQDHAPWVPVYLSIVGILAIASFCSGPGGIPFI 310  
QY 437 LTGEFFQOSORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFATICTGAIYLY 496  
DB 311 LTGEFFQOSORPAAFTIAGTVNLSNFAVGLLPFFIQKSLDTCFLVFATICTGAIYLY 370  
QY 497 FVLPEKNTYAEISQAF 514  
DB 371 FVLPEKNTYAEISQAF 388

RESULT 13  
US-10-188-186-88  
; Sequence 88, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05

; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-15  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 90  
; LENGTH: 422  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-188-186-90

Query Match 69.5%; Score 2017; DB 12; Length 422;  
Best Local Similarity 99.5%; Pred No. 1.8e-179;  
Matches 395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 137 RKHTLANNNGFAISALLMACSLQAGAFEMLVGRFIMGIDGGVALSVLPMYLSISPKE 196  
DB 11 RKHTLANNNGFAISALLMACSLQAGAFEMLVGRFIMGIDGGVALSVLPMYLSISPKE 70  
  
QY 197 IRGSLGQVTAIPICIGVTGQLLGLPELLGKSTWYLFVGVVPAVVQVLLSLPFLDPS 256  
DB 71 IRGSLGQVTAIPICIGVTGQLLGLPELLGKSTWYLFVGVVPAVVQVLLSLPFLDPS 130  
  
QY 257 RYLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVQSRIRLSVLELLRAPYVRWQ 316  
DB 131 RYLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVQSRIRLSVLELLRAPYVRWQ 190  
  
QY 317 VVTVTWVACVQLCGNLAIWFTNSIFGKAGIPPAKIPYVTLSTGDIETLAASFGLVIE 376  
DB 191 VVTVTWVACVQLCGNLAIWFTNSIFGKAGIPPAKIPYATLSTGDIETLAASFGLVIE 250  
  
QY 377 HLGRRPLIGGFLMGLFFGTITITLTLODHAPVWPYLSIVGILAIASFCSGPGGIPFI 436  
DB 251 HLGRRPLIGGFLMGLFFGTITITLTLODHAPVWPYLSIVGILAIASFCSGPGGIPFI 310  
  
QY 437 LTGEFFQSORPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTYCFLVFATCITGAIYLY 496  
DB 311 LTGEFFQSORPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTYCFLVFATCITGAIYLY 370  
  
QY 497 FVLPEKRTVAEISQAFSKRNKAYPPEKIDSATVD 533  
DB 371 FVLPEKRTVAEISQAFSKRNKAYPPEKIDSATVD 407

RESULT 10  
US-10-168-651-6  
; Sequence 6, Application US/10168651  
; Publication No. US20030171275A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: BURFORD, Neil  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: LU, Dyung Aina M.

; APPLICANT: YANG, Junming  
; APPLICANT: REDDY, Roopa  
; APPLICANT: LAL, Preeti  
; APPLICANT: HILLMAN, Jennifer L.  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: YUE, Henry  
; APPLICANT: NGUYEN, Dannie B.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: GANDHI, Ameen R.  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: KHAN, Farrah A.  
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS  
; FILE REFERENCE: PI-0005 PCT  
; CURRENT APPLICATION NUMBER: US/10/168,651  
; CURRENT FILING DATE: 2002-06-21  
; PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,75  
; 60/181,625  
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;  
; 2000-02-10  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PERL Program  
; SEQ ID NO 6  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030171275A1 6618083CD1  
US-10-168-651-6

Query Match 66.8%; Score 1939; DB 14; Length 416;  
Best Local Similarity 97.9%; Pred. No. 3.4e-172;  
Matches 380; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
QY 47 RRRKDWSCSLNVLASLAGAFSGSFLYGYNLVSVNAPPTPKAFYNESWERRHGRPDPTL 106  
DB 18 KKKLDWSCSLNVLASLAGAFSGSFLYGYNLVSVNAPPTPKAFYNESWERRHGRPDPTL 77  
  
QY 107 TLLWSVTVSIFALGVLGVLIVVMICKVLRKHTLLANNNGFAISALLMACSLQAGAFEM 166  
DB 78 TLLWSVTVSIFALGVLGVLIVVMICKVLRKHTLLANNNGFAISALLMACSLQAGAFEM 137  
  
QY 167 LIVGRITMGIDGGVALSVLPMYLSISPKEIRGSLGQVTAIFICIGVFTGQLLGLPELLG 226  
DB 138 LIVGRITMGIDGGVALSVLPMYLSISPKEIRGSLGQVTAIFICIGVFTGQLLGLPELLG 197  
  
QY 227 KESTWVPLFGVIVPAVVQVLLSLPFLDPSRVLLEKNEARAVKAFOTFLGKADVSQEV 286  
DB 198 KESTWVPLFGVIVPAVVQVLLSLPFLDPSRVLLEKNEARAVKAFOTFLGKADVSQEV 257  
  
QY 287 EYVLASHVQSRIRLSVLELLRAPYVRWQVTVITMACYQLCGNLAIWFTNSIFGKA 346  
DB 258 EYVLASHVQSRIRLSVLELLRAPYVRWQVTVITMACYQLCGNLAIWFTNSIFGKA 317  
  
QY 347 GIPPAKIPYVTLSTGDIETLAASFGLVIEHLGRRPLIGGFLMGLFFGTITITLTLOD 406  
DB 318 GIPPAKIPYVTLSTGDIETLAASFGLVIEHLGRRPLIGGFLMGLFFGTITITLTLOD 377  
  
QY 407 HAPWVPLSVIGLAIASFCSGPGGIP 434  
DB 378 HAPWVPLSVIGLAIASFCSGPGVFP 405

RESULT 11  
US-10-188-186-86  
; Sequence 86, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02

```

; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 78
; LENGTH: 565
; TYPE: PR1
; ORGANISM: Homo sapiens
US-10-188-186-78

Query Match      80.5%; Score 2337.5; DB 12; Length 565;
Best Local Similarity 77.3%; Pred. No. 3.1e-209;
Matches 474; Conservative 0; Mismatches 4; Indels 135; Gaps 2;

QY 1 MARKONRNSKELGLVPTDDTSHAGPPGCRALLCEDHLRSGVPGRRRK----- 50
Db 1 MARKONRNSKELGLVPTDDTSHAGPPGCRALLCEDHLRSGVPGRRRKQPLRSTSSAA 60
QY 51----- 50
Db 61 GSSTTYVASAAIKIPFHRWASLAPEKSSHLRLLOLRLGLEVKLISLGRKQSSISWL 120
QY 51-----DWSCSLLVASLACAFGSSFLYGNLSVNNAPTPIYKAFYNESWERRHGR 100
Db 121 ALETNIGRWDSCSLLVASLACAFGSSFLYGNLSVNNAPT----- 163
QY 101 IDPDTLLWSVTVSIFAIGLVTGTLIVRMIGKVLGRKHTLLANNGFAISALLMACSLQ 160
Db 164-----HTLLANNGFAISALLMACSLQ 185
QY 161 AGAFEMLIIVGRFINGIDGVALSVLPMYLSLSPKEIRGSLGQVTAIFICIGVFTGQLLG 220
Db 186 AGAFEMLIIVGRFINGIDGVALSVLPMYLSLSPKEIRGSLGQVTAIFICIGVFTGQLLG 245
QY 221 LPELLGKSTWPLYFGVIVVPAVVQLLSLPELPDSPRYLLLEKHNARAVKAFOTFLGKA 280
Db 246 LPELLGKSTWPLYFGVIVVPAVVQLLSLPELPDSPRYLLLEKHNARAVKAFOTFLGKA 305
QY 281 DVSEVEVLAEASHVORSIRLSVLELLRAPYRVQVTVIVTMACYOLCGNINAIWFYTN 340
Db 306 DVSEVEVLAEASHVORSIRLSVLELLRAPYRVQVTVIVTMACYOLCGNINAIWFYTN 365
QY 341 SIFKAGIPPAKIPYVTLSTGGIETLAAVFSGLVIEHLGRRLPLLIGFGLMGLFFGTULTI 400
Db 366 SIFKAGIPPAKIPYVTLSTGGIETLAAVFSGLVIEHLGRRLPLLIGFGLMGLFFGALT 425
QY 401 TLTLODHAPVPVLSIVGILAIIASFCSPGGIPILTGEFFQSQORPAAFIAGTVNWL 460
Db 426 TLTLODHAPVPVLSIVGILAIIASFCSPGGIPILTGEFFQSQORPAAFIAGTVNWL 485
QY 461 SNFVAGLLFPFIOKSLDTCFLVFATICTGAIYLYFVLPTKNTYAEISQAFSKRNKA 520
Db 486 SNFVAGLLFPFIOKSLDTCFLVFATICTGAIYLYFVLPTKNTYAEISQAFSKRNKA 545
QY 521 YPPEKIDSAVTD 533
Db 546 YPPEKIDSAVTD 558

RESULT 9
US-10-188-186-90
; Sequence 90, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188.186

```

; LENGTH: 472
; TYPE: PRT
; ORGANISM: Human
US-10-281-319-4

Query Match 82.5%; Score 2396; DB 14; Length 472;
Best Local Similarity 99.6%; Pred. No. 8.3e-215;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 55 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDDPDTLLLSWTV 114
Db 1 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDDPDTLLLSWTV 60

QY 115 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIM 174
Db 61 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIM 120

QY 175 GIDGGVALSVLPMYLSLSPKESIRGSLQGVTAIFICIGVFTGQLLGLPELLGKSTWPLY 234
Db 121 GIDGGVALSVLPMYLSLSPKESIRGSLQGVTAIFICIGVFTGQLLGLPELLGKSTWPLY 180

QY 235 FGVIVPAVQVLLSLPFLPDSPRYLLEKNEARAVKAFQFTLKGADVSQVEEVLAEASH 294
Db 181 FGVIVPAVQVLLSLPFLPDSPRYLLEKNEARAVKAFQFTLKGADVSQVEEVLAEASH 240

QY 295 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIWFYTNISFGKAGIPPAKIP 354
Db 241 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIWFYTNISFGKAGIPPAKIP 300

QY 355 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFLMGLFFGTTLTITLQDHAPWPVYL 414
Db 301 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFLMGLFFGTTLTITLQDHAPWPVYL 360

QY 415 SIVGILAIASCSGPGGIPILTGEFFQOSORPAAFIAGTVNWSNFAVGLLFPFIQK 474
Db 361 SIVGILAIASCSGPGGIPILTGEFFQOSORPAAFIAGTVNWSNFAVGLLFPFIQK 420

QY 475 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526
Db 421 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 472

RESULT 7
US-10-188-186-80
; Sequence 80, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.

; LENGTH: 472
; TYPE: PRT
; ORGANISM: Human
US-09-822-863-4

Query Match 82.5%; Score 2396; DB 9; Length 472;
Best Local Similarity 99.6%; Pred. No. 8.3e-215;
Matches 470; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 55 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDDPDTLLLSWTV 114
Db 1 SLLVASLAGAGSSFLGYNLSVNVNAPTPYIKAFYNESWERRHGRPIDDPDTLLLSWTV 60

QY 115 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIM 174
Db 61 SIFAIGGLVGLTIIVKMGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIIVGRFIM 120

QY 175 GIDGGVALSVLPMYLSLSPKESIRGSLQGVTAIFICIGVFTGQLLGLPELLGKSTWPLY 234
Db 121 GIDGGVALSVLPMYLSLSPKESIRGSLQGVTAIFICIGVFTGQLLGLPELLGKSTWPLY 180

QY 235 FGVIVPAVQVLLSLPFLPDSPRYLLEKNEARAVKAFQFTLKGADVSQVEEVLAEASH 294
Db 181 FGVIVPAVQVLLSLPFLPDSPRYLLEKNEARAVKAFQFTLKGADVSQVEEVLAEASH 240

QY 295 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIWFYTNISFGKAGIPPAKIP 354
Db 241 VORSIRLVSVLELLRAPYVRQVTVIVTMACYQLCGLNAIWFYTNISFGKAGIPPAKIP 300

QY 355 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFLMGLFFGTTLTITLQDHAPWPVYL 414
Db 301 YVTLSTGGIETLAAVFSGLVIEHLGRRELLIGGFLMGLFFGTTLTITLQDHAPWPVYL 360

QY 415 SIVGILAIASCSGPGGIPILTGEFFQOSORPAAFIAGTVNWSNFAVGLLFPFIQK 474
Db 361 SIVGILAIASCSGPGGIPILTGEFFQOSORPAAFIAGTVNWSNFAVGLLFPFIQK 420

QY 475 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526
Db 421 SLDTYCFVLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 472

RESULT 6
US-10-281-319-4
; Sequence 4, Application US/10281319
; Publication No. US20030138820A1
; GENERAL INFORMATION:
; APPLICANT: YAN, Chunhua et al.
; TITLE OF INVENTION: ISOLATED HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN TRANSPORTER PROTEINS,
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL001002CON
; CURRENT APPLICATION NUMBER: US/10/281,319
; CURRENT FILING DATE: 2002-10-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4



TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME

FILE REFERENCE: 21402-397C  
CURRENT APPLICATION NUMBER: US/10/188,186  
CURRENT FILING DATE: 2002-07-02  
PRIOR APPLICATION NUMBER: 60/303046  
PRIOR FILING DATE: 2001-07-05  
PRIOR APPLICATION NUMBER: 60/360814  
PRIOR FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: 60/303828  
PRIOR FILING DATE: 2001-09-07  
PRIOR APPLICATION NUMBER: 60/323380  
PRIOR FILING DATE: 2001-09-19  
PRIOR APPLICATION NUMBER: 60/361133  
PRIOR FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: 60/304016  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: 60/304502  
PRIOR FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: 60/305262  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: 60/373881  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: 60/305673  
PRIOR FILING DATE: 2001-07-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: Custom  
SEQ ID NO 76  
LENGTH: 507  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-188-186-76

Query Match 93.8%; Score 2723; DB 15; Length 537;  
Best Local Similarity 99.4%; Pred. No. 2.8e-245;  
Matches 531; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWCSLLVAS 60  
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWCSLLVAS 60  
QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHRRPDPDPTLLTLLSVTVSIFAIG 120  
DB 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHRRPDPDPTLLTLLSVTVSIFAIG 120  
QY 121 GLVGTLLIVKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180  
DB 121 GLVGTLLIVKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180  
QY 181 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 240  
DB 181 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 240  
QY 241 PAVVQLLSLPFLPDSPRYLLEKHNARAVKAFQTFGLKADYSQVEEVLAEASHVQSRIR 300  
DB 241 PAVVQLLSLPFLPDSPRYLLEKHNARAVKAFQTFGLKADYSQVEEVLAEASHVQSRIR 300  
QY 301 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFTYNSIPGKAGIPPAKIPYVTLST 360  
DB 301 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFTYNSIPGKAGIPPAKIPYVTLST 360  
QY 361 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPVYLSIVGIL 420  
DB 361 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPVYLSIVGIL 420  
QY 421 AIIASFCSGPGGIPFLTGTGFFQOSQRPAAFIAGTVNWSNFAVGLLPFTQKSLDTYC 480  
DB 421 AIIASFCSGPGGIPFLTGTGFFQOSQRPAAFIAGTVNWSNFAVGLLPFTQKSLDTYC 480  
QY 481 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 534  
DB 481 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 534  
RESULT 4  
US-10-188-186-76  
Sequence 76, Application US/10188186  
Publication No. US20040029789A1  
GENERAL INFORMATION:  
APPLICANT: Anderson et al.

FILE REFERENCE: 21402-397C  
CURRENT APPLICATION NUMBER: US/10/188,186  
CURRENT FILING DATE: 2002-07-02  
PRIOR APPLICATION NUMBER: 60/303046  
PRIOR FILING DATE: 2001-07-05  
PRIOR APPLICATION NUMBER: 60/360814  
PRIOR FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: 60/303828  
PRIOR FILING DATE: 2001-09-07  
PRIOR APPLICATION NUMBER: 60/323380  
PRIOR FILING DATE: 2001-09-19  
PRIOR APPLICATION NUMBER: 60/361133  
PRIOR FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: 60/304016  
PRIOR FILING DATE: 2001-07-09  
PRIOR APPLICATION NUMBER: 60/304502  
PRIOR FILING DATE: 2001-07-11  
PRIOR APPLICATION NUMBER: 60/305262  
PRIOR FILING DATE: 2001-07-13  
PRIOR APPLICATION NUMBER: 60/373881  
PRIOR FILING DATE: 2002-04-19  
PRIOR APPLICATION NUMBER: 60/305673  
PRIOR FILING DATE: 2001-07-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 368  
SOFTWARE: Custom  
SEQ ID NO 76  
LENGTH: 507  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-188-186-76  
Query Match 87.1%; Score 2527.5; DB 12; Length 507;  
Best Local Similarity 93.4%; Pred. No. 4.8e-227;  
Matches 498; Conservative 0; Mismatches 2; Indels 33; Gaps 1;  
QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWCSLLVAS 60  
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWCSLLVAS 60  
QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHRRPDPDPTLLTLLSVTVSIFAIG 120  
DB 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHRRPDPDPTLLTLLSVTVSIFAIG 120  
QY 121 GLVGTLLIVKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180  
DB 121 GLVGTLLIVKMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIIVGRFIMGIDGV 180  
QY 181 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 240  
DB 181 ALSVLPWYLSISPKIRSGSLGQVTAIFICIGVFTGQLLGPPELLGKSTWYPLFGVIV 240  
QY 241 PAVVQLLSLPFLPDSPRYLLEKHNARAVKAFQTFGLKADYSQVEEVLAEASHVQSRIR 300  
DB 241 PAVVQLLSLPFLPDSPRYLLEKHNARAVKAFQTFGLKADYSQVEEVLAEASHVQSRIR 300  
QY 301 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFTYNSIPGKAGIPPAKIPYVTLST 360  
DB 301 LVSVELELRAPYVRWQVTVITMACYQLCGLNAINWFTYNSIPGKAGIPPAKIPYVTLST 360  
QY 361 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPVYLSIVGIL 420  
DB 361 GGIETLAAPVSGLVIEHGRRLPILGFGMLGFFGTLITLTLQDHAPWPVYLSIVGIL 420  
QY 421 AIIASFCSGPGGIPFLTGTGFFQOSQRPAAFIAGTVNWSNFAVGLLPFTQKSLDTYC 480  
DB 421 AIIASFCSGPGGIPFLTGTGFFQOSQRPAAFIAGTVNWSNFAVGLLPFTQKSLDTYC 480  
QY 481 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 533  
DB 481 FLVPATICITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 500

US-10-276-774-2280

ORGANISM: Homo sapiens

Query Match 94.3%; Score 2737; DB 12; Length 558;  
Best Local Similarity 99.8%; Pred. No. 1.5e-246;  
Matches 532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 60  
DB 19 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 78  
QY 61 LAGAFGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 120  
DB 79 LAGAFGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 138  
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNFPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGV 180  
DB 139 GLVGTLLIVKMGKVLGRKHTLLANNFPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGV 198  
QY 181 ALSVLPWYLSSEISPKIRGSLGQVTAIFICIGVFTGQLGLPELLGKSTWPLYFGVIVV 240  
DB 199 ALSVLPWYLSSEISPKIRGSLGQVTAIFICIGVFTGQLGLPELLGKSTWPLYFGVIVV 258  
QY 241 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 300  
DB 259 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 318  
QY 301 LVSUVEILLRAPPYVRQVTVVITWACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 360  
DB 319 LVSUVEILLRAPPYVRQVTVVITWACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 378  
QY 361 GGIEETLAAVFSGLVIEHLGRPLIIGGFLMGLPFGTITITLQDHAPWVPLYSIVGIL 420  
DB 379 GGIEETLAAVFSGLVIEHLGRPLIIGGFLMGLPFGTITITLQDHAPWVPLYSIVGIL 438  
QY 421 AIIASFSGPGGIPFILTGEFFQSQORPAAFIAGTVNLSNFAVGLLFPFIQKSLDTC 480  
DB 439 AIIASFSGPGGIPFILTGEFFQSQORPAAFIAGTVNLSNFAVGLLFPFIQKSLDTC 498  
QY 481 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSATVD 533  
DB 499 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSATVD 551

RESULT 3

US-10-297-022-25

Sequence 25, Application US/10297022  
Publication No. US20030216310A1

GENERAL INFORMATION:  
APPLICANT: INCYTE GENOMICS, INC.  
APPLICANT: THORNTON, Michael  
APPLICANT: WALIA, Narinder K.  
APPLICANT: YUE, Henry  
APPLICANT: NGUYEN, Dannie B.  
APPLICANT: LAL, Preeti  
APPLICANT: GANDHI, Ameena R.  
APPLICANT: TRIBOULEY, Catherine M.  
APPLICANT: YAO, Monique G.  
APPLICANT: RAMKUMAR, Jayalaxmi  
APPLICANT: AU-YOUNG, Janice  
APPLICANT: LU, Yan  
APPLICANT: TANG, Y. Tom  
APPLICANT: AZIMZAI, Valda  
APPLICANT: BRUNS, Christopher M.  
APPLICANT: GRIFFIN, Jennifer A.  
APPLICANT: YANG, Junming  
APPLICANT: BAUGHN, Mariah R.  
APPLICANT: SANJANWALA, Madhu S.  
APPLICANT: RAUMANN, Brigitte E.  
APPLICANT: LEE, Ernestine A.  
APPLICANT: HAFALIA, April  
APPLICANT: GREENE, Barrie D.  
APPLICANT: KHAN, Farrah A.

US-09-981-947A-2

LENGTH: 563 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 2;

Query Match 100.0%; Score 2903; DB 9; Length 563;  
Best Local Similarity 100.0%; Pred. No. 4.7e-262;  
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 60  
DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSVGPGRRRKDWSCSLIVAS 60  
QY 61 LAGAFGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 120  
DB 61 LAGAFGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 120  
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNFPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGV 180  
DB 121 GLVGTLLIVKMGKVLGRKHTLLANNFPAISAAALLMACSLOAGAFEMLIIVGRFIMGIDGV 180  
QY 181 ALSVLPWYLSSEISPKIRGSLGQVTAIFICIGVFTGQLGLPELLGKSTWPLYFGVIVV 240  
DB 181 ALSVLPWYLSSEISPKIRGSLGQVTAIFICIGVFTGQLGLPELLGKSTWPLYFGVIVV 240  
QY 241 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 300  
DB 241 PAVVQLLSLPLDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVVEVLAESHVORSIR 300  
QY 301 LVSUVEILLRAPPYVRQVTVVITWACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 360  
DB 301 LVSUVEILLRAPPYVRQVTVVITWACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 360  
QY 361 GGIEETLAAVFSGLVIEHLGRPLIIGGFLMGLPFGTITITLQDHAPWVPLYSIVGIL 420  
DB 361 GGIEETLAAVFSGLVIEHLGRPLIIGGFLMGLPFGTITITLQDHAPWVPLYSIVGIL 420  
QY 421 AIIASFSGPGGIPFILTGEFFQSQORPAAFIAGTVNLSNFAVGLLFPFIQKSLDTC 480  
DB 421 AIIASFSGPGGIPFILTGEFFQSQORPAAFIAGTVNLSNFAVGLLFPFIQKSLDTC 480  
QY 481 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSATVDAPASSPF 540  
DB 481 FLVFTATCITGAIYLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSATVDAPASSPF 540  
QY 541 TTPNTAWTQAAATTTATKKEHPL 563  
DB 541 TTPNTAWTQAAATTTATKKEHPL 563

RESULT 2

US-10-276-774-2280

Sequence 2280, Application US/10276774  
Publication No. US20040053245A1

GENERAL INFORMATION:  
APPLICANT: Hyseq, Inc.  
APPLICANT: Tang, Y. Tom et al  
TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides  
FILE REFERENCE: 21272-030  
CURRENT APPLICATION NUMBER: US/10/276,774  
CURRENT FILING DATE: 2002-11-18  
PRIOR APPLICATION NUMBER: 09/560,875  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 09/496,914  
PRIOR FILING DATE: 2000-02-03  
NUMBER OF SEQ ID NOS: 2700  
SOFTWARE: Custom  
SEQ ID NO 2280  
LENGTH: 558  
TYPE: PRT

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:18:33 ; Search time 46 Seconds  
(without alignments)  
3214.404 Million cell updates/sec

Title: US-09-981-947B-2

Perfect score: 2903

Sequence: 1 MARKQNRNSKGLVPLTDD.....NTAWTQAAATTTATKKEHPL 563

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1071772 seqs, 262633353 residues

Total number of hits satisfying chosen parameters: 1071772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep:\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep:\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep:\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep:\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep:\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep:\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep:\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep:\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep:\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep:\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep:\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep:\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep:\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep:\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep:\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep:\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep:\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2903	100.0	563	9	US-09-981-947A-2
2	2737	94.3	558	12	US-10-276-774-2280
3	2723	93.8	537	15	US-10-297-022-25
4	2527.5	87.1	507	12	US-10-188-186-76
5	2396	82.5	472	9	US-09-822-863-4
6	2396	82.5	472	14	US-10-281-319-4
7	2341	80.6	480	12	US-10-188-186-80
8	2337.5	80.5	565	12	US-10-188-186-78
9	2017	69.5	422	12	US-10-188-186-90
10	1939	66.8	416	14	US-10-168-651-6
11	1929	66.4	396	12	US-10-188-186-86
12	1920	66.1	396	12	US-10-188-186-84
13	1911	65.8	396	12	US-10-188-186-88
14	1721	59.3	362	12	US-10-188-186-82
15	1122.5	38.7	501	15	US-10-295-027-308

16	1119.5	38.6	512	12	US-10-332-447-7	Sequence 7, Appli
17	1112.5	38.3	502	12	US-10-029-020-65	Sequence 65, Appl
18	1109.5	38.2	502	12	US-10-029-020-66	Sequence 66, Appl
19	1108	38.2	500	9	US-09-981-947A-7	Sequence 7, Appli
20	1103.5	38.0	502	12	US-10-029-020-67	Sequence 67, Appl
21	1103.5	38.0	524	12	US-10-029-020-64	Sequence 64, Appl
22	1099.5	37.9	501	12	US-10-029-020-68	Sequence 68, Appl
23	1081	37.2	493	9	US-09-981-947A-10	Sequence 10, Appl
24	1032.5	35.6	481	15	US-10-116-275-345	Sequence 345, App
25	1000.5	34.5	499	12	US-10-332-447-15	Sequence 15, Appl
26	997.5	34.4	483	12	US-10-029-020-24	Sequence 24, Appl
27	979	33.7	524	13	US-10-094-059-2	Sequence 2, Appli
28	936.5	32.3	492	15	US-10-295-027-1248	Sequence 1248, Ap
29	936	32.2	471	9	US-09-778-927A-47	Sequence 47, Appl
30	935	32.2	501	9	US-09-778-927A-48	Sequence 48, Appl
31	934.5	32.2	480	14	US-10-176-847-76	Sequence 76, Appl
32	931.5	32.1	492	14	US-10-328-198-3	Sequence 3, Appli
33	897.5	30.9	516	9	US-09-822-863-2	Sequence 2, Appli
34	897.5	30.9	516	14	US-10-281-319-2	Sequence 2, Appli
35	883	30.4	455	9	US-09-778-927A-50	Sequence 50, Appl
36	879	30.3	471	9	US-09-778-927A-49	Sequence 49, Appl
37	874	30.1	503	9	US-09-981-947A-6	Sequence 6, Appli
38	866	29.8	494	9	US-09-981-947A-5	Sequence 5, Appli
39	842	29.0	441	9	US-09-778-927A-51	Sequence 51, Appl
40	830	28.6	496	15	US-10-341-434-89	Sequence 89, Appl
41	773	26.6	522	15	US-10-099-322-124	Sequence 124, App
42	773	26.6	522	15	US-10-044-564-124	Sequence 124, App
43	772	26.6	522	15	US-10-099-322-125	Sequence 125, App
44	772	26.6	522	15	US-10-044-564-125	Sequence 125, App
45	767.5	26.4	523	15	US-10-099-322-122	Sequence 122, App

ALIGNMENTS

RESULT 1  
; Sequence 2, Application US/09981947A  
; Patent No. US20020164578A1  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/981.947A  
; FILING DATE: 18-Oct-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031.392  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:

Db 10 LIPASIAIGSFQFYNTGVINAPEAIKDFLNTLEERSETPSPSSVLLTSLMSLSVAI 69  
 QY 117 FAIGGLVGLIYKMLCKVLGRKHTLLANNNGFAISAAALLMACSLQAGAPEMLIVGRFTWGI 176  
 Db 70 FSVGGMIGSFVGLFVNRFRNSMLIVNLAIAGGLMGFCCKIAESVEMILGRLLIIGL 129  
 QY 177 DGGVALSVLPMYLSISPKIRSGSLGQVTAIFCIGVFTGQLLGLPELLGKKESTWPLYFG 236  
 Db 130 FGLCTGFPWYIGISPTALRGAFTLNQLGIVIGILVAQIFGLKVLGTEDLWPLJLG 189  
 QY 237 VIVPRAVOLLSLPFLPDSPRYLLEKNEARAVKAFQTFGLKADVSQOEVEEVLAEHVQ 296  
 Db 190 FTILPAIIQCAALPFCPSPRELLINRKEEKAKEILQRLWGTEDVAQDIQEMKDESIRM 249  
 QY 297 RSIRLVSVLELARAPVPRWQVVTIVTMACVOLCGLNAIWFVWNSIFGKAGIPPAKIPYV 356  
 Db 250 SOEKQVTVLELFRAPNYQPIIISIMLOLSQSLGINAVFYISTGIFKADGV--QEPVYA 307  
 QY 357 TLSTGGIETLAAVFSGLVIEHLGRPLLIIGFGLMGLEFGTLITITLQDHAPWVPYLSI 416  
 Db 308 TIGAGVWNTIFTVSVFELVERAGRTLHLIGLGGWAFCSILMTISLLKONYSWMSFICI 367  
 QY 417 VGIILAIASFCSGGPGGIPFILTGEFFQOSQRPAPFIAGTVNWLNSFVAGLLFPFIQKSL 476  
 Db 368 GAILVFAFFETIGPGPIFWFIVAELFGQPRPAAAMAVAGCSNWTNLFVLGILLFPFSATYL 427  
 QY 477 DTYCFLVEATICITGAILVYFVLPETKNTYAEISQAF 514  
 Db 428 GAYFIVFTVFLVIFWVFTFFKVPETGRTEETRAF 465

Search completed: April 6, 2004, 10:19:34  
 Job time : 24 secs



Qy 411 VYLSIVGILAIASFCSPGIPILTGFFQSQRPAAPIIAGTWNWLSNFAVGLLFP 470  
 Db 380 MSYVSIVAFEGVAFPEIGPGPIFWIVAEIIFSQQPRPAAVAVAGFSNWTSTIGMGFP 439  
 Qy 471 FIQKSLDTCFLVFNATICITGAIVLYFVLPEKNTVAELISQAFSKRNKAYPPEEK 526  
 Db 440 YVAEAWGPYVLLFAVLLGFFITFLRVPETRTDQISAAFRHPSLLEQEVK 495

RESULT 13  
 US-09-299-549-6  
 ; Sequence 6, Application US/09299549  
 ; Patent No. 6136547  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tartaglia, Louis A.  
 ; APPLICANT: Weng, Xun  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
 ; TITLE OF INVENTION: GLUTEX AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: Windows95  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/299,549  
 ; FILING DATE: 26-APR-1999  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 09/031,392  
 ; FILING DATE: 26-FEB-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Meiklejohn, Ph.D., Anita L.  
 ; REGISTRATION NUMBER: 35,283  
 ; REFERENCE/DOCKET NUMBER: 07334/072002  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617/542-5070  
 ; TELEFAX: 617/542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 509 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-299-549-6

Query Match 30.1%; Score 874; DB 3; Length 509;  
 Best Local Similarity 38.7%; Pred. No. 3.4e-78;  
 Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;

Qy 57 LVASLAGFGSSFLYGNLSVNVNAPTPYIKAFYNESWERRHG-----RPIDPDTLLWSV 112  
 Db 24 LVAVFSAVLGSLQFGYGNIGVINAPQKVIQSYNETWLGKRGPGPSSIPPGTLLTVAL 83  
 Qy 113 TVSIFAIGGLVGTLLVKGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIYGRF 172  
 Db 84 SVAIFSVMGMISSFLIGISQWLGKRAMLVNVLAVLGSLGSLMGLANAAASYEMLIIGRF 143  
 Qy 173 IMGIDGGVALSVLPMYLSLSEIPKIRGSLGQVTAIFTCIGVFTGQLLGLPGLLKESTWP 232  
 Db 144 LGAGSGLTSLGVLPMYVGEIAPHRLAGALGTNLQALVIGILIAQVGLSLLGTASLWP 203  
 Qy 233 YLFGVIVPVAVQLLSPLPDSPRVYLLEKNEARAVKAFQTFGLKADVQSEVEVLAE 292  
 Db 204 LLGLTLPALLQVLLPFCPSPRYLIQNLGEPARKSLKRLTGTWADVSGVLAELKDE 263

Db 429 CGPYVFIIFTVLLVLFIRTYKVPETKGRITDEIASGFRQGGASQDKT--PEE 481

RESULT 12  
 US-09-031-392-6  
 ; Sequence 6, Application US/09031392  
 ; Patent No. 5942398  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tartaglia, Louis A.  
 ; APPLICANT: Weng, Xun  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES  
 ; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: Windows95  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/031,392  
 ; FILING DATE: 26-FEB-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Meiklejohn, Ph.D., Anita L.  
 ; REGISTRATION NUMBER: 35,283  
 ; REFERENCE/DOCKET NUMBER: 07334/072001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617/542-5070  
 ; TELEFAX: 617/542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 6:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 509 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-031-392-6

Query Match 30.1%; Score 874; DB 2; Length 509;  
 Best Local Similarity 38.7%; Pred. No. 3.4e-78;  
 Matches 184; Conservative 99; Mismatches 183; Indels 10; Gaps 5;

Qy 57 LVASLAGFGSSFLYGNLSVNVNAPTPYIKAFYNESWERRHG-----RPIDPDTLLWSV 112  
 Db 24 LVAVFSAVLGSLQFGYGNIGVINAPQKVIQSYNETWLGKRGPGPSSIPPGTLLTVAL 83  
 Qy 113 TVSIFAIGGLVGTLLVKGKVLGRKHTLLANNNGFAISALLMACSLQAGAFEMLIYGRF 172  
 Db 84 SVAIFSVMGMISSFLIGISQWLGKRAMLVNVLAVLGSLGSLMGLANAAASYEMLIIGRF 143  
 Qy 173 IMGIDGGVALSVLPMYLSLSEIPKIRGSLGQVTAIFTCIGVFTGQLLGLPGLLKESTWP 232  
 Db 144 LGAGSGLTSLGVLPMYVGEIAPHRLAGALGTNLQALVIGILIAQVGLSLLGTASLWP 203  
 Qy 233 YLFGVIVPVAVQLLSPLPDSPRVYLLEKNEARAVKAFQTFGLKADVQSEVEVLAE 292  
 Db 204 LLGLTLPALLQVLLPFCPSPRYLIQNLGEPARKSLKRLTGTWADVSGVLAELKDE 263  
 Qy 293 SHVORSTRVLSVLELLRAPVVRQVTVITWACVQLCLNAINWFTVNSIFGKAGI-PPA 351  
 Db 264 KRKLERERPLSLQLGSRTHRQPLIIAVVQLSGLNINAVFYYSIFETAGVGQPA 323  
 Qy 352 KIPYVTLSTGISTLAASVGLVIEHLGRPL-LIGGFLGMLGFTGTITITLTLDHAPW 410  
 Db 324 ---YATIGAGVNVTVTLVSLVILVERGRRTLHLGLAGMCCAI-LMTVALLLLIERYPA 379

STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10112-0228  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/16126  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/355,844  
FILING DATE: 14-DEC-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Tang, Henry Y.S.  
REGISTRATION NUMBER: 29,705  
REFERENCE/DOCKET NUMBER: A29927-50/29910  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-408-2586  
TELEFAX: 212-765-2519  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 492 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE: Human  
ORGANISM: Human  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..492  
OTHER INFORMATION: Facilitative glucose transporter  
OTHER INFORMATION: Glut1 protein  
PCT-US95-16126-3  
Query Match 31.7%; Score 919.5; DB 5; Length 492;  
Best Local Similarity 39.2%; Pred. No. 9.4e-83;  
Matches 186; Conservative 98; Mismatches 180; Indels 11; Gaps 5;  
QY 57 LVASLAGAGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLTLLMSVTSI 116  
DB 12 LMLAVGAVLGLSQFGYNTGVINAPQKVIEFYNTQWVHRYGESILPTLTLLWSLSVAI 71  
QY 117 FAIGGLVGLTLIVKMGKVLGRKHTLLANNGFAISALLMACSLQAGAFEMLIIVGRFIMG 176  
DB 72 FSVGGMIGSFVGLFVNRFGRRNSMLMNLALLAFVSALVLMGFSKLGKSFEMILIGRFIIV 131  
QY 177 DGGVALSVLPMYLSIEISPKIRSGLVQVTAIFICIGVFTGQLGLPELKGKSTWPLYFG 236  
DB 132 YCGLTTGFPVMYGVESPTAFRGALGTLHQLGIVVGLIIAQVFGLDISMGKDLWPLLS 191  
QY 237 VIVPVAVQLLSLPLDPSRYLLEKNEARAVKAFQTLGKADVSQVEEVLAEHVQ 296  
DB 192 IIFIPALLQCVIPFCPSPRFLINRNEENRAKSVLKLGRGTADVTHDLQEMKEESRQM 251  
QY 297 RSIRLVSVLELLRAPYVRQVTVITMACYQLCGLNATWFTNSIFGKAGIPPAKIPV 356  
DB 252 MREKVTILELFRSPAYRQPIILIAVLQLSQQLSGINAVFYSTSIPEKAGV--QQPYVA 309  
QY 357 TLSTGGIETLAAVFSGLVIEHLGRRL-LIGGFGMLGFFGTTLITLTLDHAPWPVYLS 415  
DB 310 TIGSGIVNTAFTVSLFVVERAGRTLHLIGLAGMAGQAI-LMTIALALLEQLPWSYLS 368  
QY 416 IVGILAIISFCSPGGIPFLTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQKS 475  
DB 369 IVAIFGVAFVFEVGPPIWFIVAELESQGPRAIAVAGFSNWTNFIIVGMCQFYVEQL 428  
QY 476 LDTYCFELVATICITGAIYLYFVLPETKNTYABISQAF-----SKNKAYPPEE 525

NAME: Tang, Henry Y.S.  
REGISTRATION NUMBER: 29,705  
REFERENCE/DOCKET NUMBER: A29927-50/29910  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-408-2586  
TELEFAX: 212-765-2519  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 492 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
ORIGINAL SOURCE: Human  
ORGANISM: Human  
FEATURE:  
NAME/KEY: Peptide  
LOCATION: 1..492  
OTHER INFORMATION: Facilitative glucose transporter  
OTHER INFORMATION: Glut1 protein  
US-08-355-844-3  
Query Match 31.7%; Score 919.5; DB 2; Length 492;  
Best Local Similarity 39.2%; Pred. No. 9.4e-83;  
Matches 186; Conservative 98; Mismatches 180; Indels 11; Gaps 5;  
QY 57 LVASLAGAGSSFLYGYNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLTLLMSVTSI 116  
DB 12 LMLAVGAVLGLSQFGYNTGVINAPQKVIEFYNTQWVHRYGESILPTLTLLWSLSVAI 71  
QY 117 FAIGGLVGLTLIVKMGKVLGRKHTLLANNGFAISALLMACSLQAGAFEMLIIVGRFIMG 176  
DB 72 FSVGGMIGSFVGLFVNRFGRRNSMLMNLALLAFVSALVLMGFSKLGKSFEMILIGRFIIV 131  
QY 177 DGGVALSVLPMYLSIEISPKIRSGLVQVTAIFICIGVFTGQLGLPELKGKSTWPLYFG 236  
DB 132 YCGLTTGFPVMYGVESPTAFRGALGTLHQLGIVVGLIIAQVFGLDISMGKDLWPLLS 191  
QY 237 VIVPVAVQLLSLPLDPSRYLLEKNEARAVKAFQTLGKADVSQVEEVLAEHVQ 296  
DB 192 IIFIPALLQCVIPFCPSPRFLINRNEENRAKSVLKLGRGTADVTHDLQEMKEESRQM 251  
QY 297 RSIRLVSVLELLRAPYVRQVTVITMACYQLCGLNATWFTNSIFGKAGIPPAKIPV 356  
DB 252 MREKVTILELFRSPAYRQPIILIAVLQLSQQLSGINAVFYSTSIPEKAGV--QQPYVA 309  
QY 357 TLSTGGIETLAAVFSGLVIEHLGRRL-LIGGFGMLGFFGTTLITLTLDHAPWPVYLS 415  
DB 310 TIGSGIVNTAFTVSLFVVERAGRTLHLIGLAGMAGQAI-LMTIALALLEQLPWSYLS 368  
QY 416 IVGILAIISFCSPGGIPFLTGEFFQSQORPAAFIAGTVNWSNFAVGLLFPFIQKS 475  
DB 369 IVAIFGVAFVFEVGPPIWFIVAELESQGPRAIAVAGFSNWTNFIIVGMCQFYVEQL 428  
QY 476 LDTYCFELVATICITGAIYLYFVLPETKNTYABISQAF-----SKNKAYPPEE 525  
DB 429 CGPYVFIETVLVLFIRTYFKVPTKGTDEIASGFRQGSQSDXT--PEE 481  
RESULT 11  
PCT-US95-16126-3  
Sequence 3, Application PC/TUS9516126  
GENERAL INFORMATION:  
APPLICANT: Fischberg, Jorge  
APPLICANT: Czegledy, Ferenc  
APPLICANT: Iserovich, Pavel  
APPLICANT: Li, Jun  
APPLICANT: Cheung, Min  
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN  
STRUCTURE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond



TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 10:  
US-09-610-417-10

Query Match 37.2%; Score 1081; DB 4; Length 493;  
Best local Similarity 48.5%; Pred. No. 7.8e-99;  
Matches 229; Conservative 71; Mismatches 168; Indels 4; Gaps 3;

QY 55 SLLVASLAGAGSSFLYGYNLVVNAPTPIKAFYNESWERRHGRPIDDTITLLWSVTV 114  
DB 13 TLVLAFLAALG-SFOYGYNLGVINAPQKVEAFY-ETWLGXGEXPSVPTITLLWSLSV 70  
QY 115 SIFAIGGLVGTLLIVKMIGKVLGRKHTLLANNPFAISAALLMACSLQAGAFEMLIIVGRFIM 174  
DB 71 SIFAVGGMIGSFLVGVXIGNRLGRKXAMLVNVLAIAGGLIMGLAKAXSFEMLIIGRFII 130  
QY 175 GIDGGVALSVLPMYLSPEISPEIRSGISGOVTAIFICIGVFTGQLLGLPELLGKESWPYL 234  
DB 131 GLYCLSGGVPMYVGEISPTALRGALGTNLQGLVIGILIAQVLGDLDSLLGNESLWPLL 190  
QY 235 FGIVVPAVAVOLLSPFLPDSPRYLLLEKHNKNEARAVKAFQTLGKADVSQVEEVLAESEH 294  
DB 191 LGITGVPAQLQLLLPFCPSFRYLLINKNEARAKKALQRLGTADVSQVEAEKMDER 250  
QY 295 VORSIRLVSVLELRAPYVRQWVTVVITMACYQLCGLNAINFTYNSIFGKAGIPPAKIP 354  
DB 251 XMKSEKXSVLELFRSRYRQPVIIAIVLQSLSGINAVFYSTISFEKAGV--GQPV 308  
QY 355 YVTLSTGGIETLAIVFSGLVIEHLGRPELLIGGGLMGLFGTTLTITLQDHAPWVYPL 414  
DB 309 YATIGAGVNVTVFTVSVFVERAGRRTLLHLLGGNAGCAVMTIALALLDQVPWMSYV 368  
QY 415 SIIVGILAIASFCSGPGGIPILTGEFFQOSQORPAAFIAGTVNWSNFAVGLLFPFIQK 474  
DB 369 SIIVAFGVFAFVEGPGIPFWFVAELFSQGPRAAIAVAGFSNWTNFIIVGLLFQYIAE 428  
QY 475 SLDTYCLVFLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526  
DB 429 LLGPYFVIVFAVLLLLFFITFLKVPETKGRFTDEIAAFKRNKXEQPEKE 480

RESULT 10  
US-08-355-844-3  
Sequence 3, Application US/08355844  
Patent No. 5940307  
GENERAL INFORMATION:  
APPLICANT: Fischbarg, Jorge  
APPLICANT: Czegledy, Ferenc  
APPLICANT: Iserovich, Pavel  
APPLICANT: Li, Jun  
APPLICANT: Cheung, Min  
TITLE OF INVENTION: A METHOD FOR PREDICTING PROTEIN  
TITLE OF INVENTION: STRUCTURE  
NUMBER OF SEQUENCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Brumbaugh, Graves, Donohue & Raymond  
STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: NY  
COUNTRY: USA  
ZIP: 10112-0228  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/355,844  
FILING DATE: 14-DEC-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:

QY 55 SLLVASLAGAGSSFLYGYNLVVNAPTPIKAFYNESWERRHGRPIDDTITLLWSVTV 114  
DB 13 TLVLAFLAALG-SFOYGYNLGVINAPQKVEAFY-ETWLGXGEXPSVPTITLLWSLSV 70  
QY 115 SIFAIGGLVGTLLIVKMIGKVLGRKHTLLANNPFAISAALLMACSLQAGAFEMLIIVGRFIM 174  
DB 71 SIFAVGGMIGSFLVGVXIGNRLGRKXAMLVNVLAIAGGLIMGLAKAXSFEMLIIGRFII 130  
QY 175 GIDGGVALSVLPMYLSPEISPEIRSGISGOVTAIFICIGVFTGQLLGLPELLGKESWPYL 234  
DB 131 GLYCLSGGVPMYVGEISPTALRGALGTNLQGLVIGILIAQVLGDLDSLLGNESLWPLL 190  
QY 235 FGIVVPAVAVOLLSPFLPDSPRYLLLEKHNKNEARAVKAFQTLGKADVSQVEEVLAESEH 294  
DB 191 LGITGVPAQLQLLLPFCPSFRYLLINKNEARAKKALQRLGTADVSQVEAEKMDER 250  
QY 295 VORSIRLVSVLELRAPYVRQWVTVVITMACYQLCGLNAINFTYNSIFGKAGIPPAKIP 354  
DB 251 XMKSEKXSVLELFRSRYRQPVIIAIVLQSLSGINAVFYSTISFEKAGV--GQPV 308  
QY 355 YVTLSTGGIETLAIVFSGLVIEHLGRPELLIGGGLMGLFGTTLTITLQDHAPWVYPL 414  
DB 309 YATIGAGVNVTVFTVSVFVERAGRRTLLHLLGGNAGCAVMTIALALLDQVPWMSYV 368  
QY 415 SIIVGILAIASFCSGPGGIPILTGEFFQOSQORPAAFIAGTVNWSNFAVGLLFPFIQK 474  
DB 369 SIIVAFGVFAFVEGPGIPFWFVAELFSQGPRAAIAVAGFSNWTNFIIVGLLFQYIAE 428  
QY 475 SLDTYCLVFLVATICITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEK 526  
DB 429 LLGPYFVIVFAVLLLLFFITFLKVPETKGRFTDEIAAFKRNKXEQPEKE 480

RESULT 9  
US-09-610-417-10  
Sequence 10, Application US/09610417  
Patent No. 6346374  
GENERAL INFORMATION:  
APPLICANT: Tartaglia, Louis A.  
APPLICANT: Weng, Xun  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
GLUTEX AND USES THEREOF  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Windows95  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/610,417  
FILING DATE: 05-Jul-2000  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/299,549  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Meiklejohn, Ph.D., Anita L.  
REGISTRATION NUMBER: 35,283  
REFERENCE/DOCKET NUMBER: 07334/072002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/542-5070  
TELEFAX: 617/542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 493 amino acids

QY 236 GTVVFAVQLLSLPDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVLEVAESHV 295  
DB 197 GTGVPAAQLQLLPFPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVLEVAESHV 256  
QY 296 QRSIRVSVLELLRAPPYVQVVTIVTMACYQLCGNNAIFVYTNISIFGKAGIPPAKIPY 355  
DB 257 ERAAGFISVLKFRMSLAWQLLSIIVLMGGQQLSGVNAIYYAQIYLSAGVPEEHVQY 316  
QY 356 VTLSTGGIE---TLAAVFSGLVIEHLGRPLIGGFLMGLPFGTLTITLTLQDHAPWVP 412  
DB 317 VTAGTGANVVMWTFCAVF---VVELLGRLLLLGFSICLIACCVLTAALALQDTVSWMP 373  
QY 413 YLSIVGILAIASFCSPGGIPILTGEFFQOSQORPAARIIAGTVNWLNSNEAVGLLFPPI 472  
DB 374 YLSIVGILAIASFCSPGGIPILTGEFFQOSQORPAARIIAGTVNWLNSNEAVGLLFPPI 432  
QY 473 QKSLDYCYFLVFATCITGAIYLYFVLPETKRTYAEISQAFSKRKNK---AYPPEKI 527  
DB 433 QKSLDYCYFLVFATCITGAIYLYFVLPETKRTYAEISQAFSKRKNK---AYPPEKI 490

RESULT 7

US-09-031-392-10  
; Sequence 10, Application US/09031392  
; Patent No. 5942398  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; APPLICANT: Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 493 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-031-392-10

Query Match 37.2%; Score 1081; DB 2; Length 493;  
Best Local Similarity 48.5%; Pred. No. 7.8e-99;  
Matches 229; Conservative 71; Mismatches 168; Indels 4; Gaps 3;  
QY 55 SILVASLACAFSSFLYGNSLVNAPTYIKAFYNESWERHGRPIDPDTLLMSVTV 114  
DB 13 TLVLAVLIAALG-SFOYGNLGVINAPQKVIEAFY-ETWLGRXGEXPSPTLLTLLWSLV 70  
QY 115 SIFAGGLVGTLLVMKIGVLRKHTLLANNGFATSAALLMACSLQAGAFEMLVGRFTM 174  
DB 71 SIFAVGMLGSLVGVIGNLGRKXAMVNVNVAIAGGLMGLMGLAXXAXSFEMLVGRFTI 130

QY 175 GIDGVALSVLPMYLSIPKIRSGLSQVTAIFICIGVFTGQLLGPPELLGKESTWPLY 234  
DB 131 GLYCLSGGVPMVVGISPTALRGALGTNLQIGIVIGILIAQVLGLDLSLLGNESLWPLL 190  
QY 235 FGVIWFAVQLLSLPDPSPRYLLLEKNEARAVKAFOTFLGKADVSQEVLEVAESHV 294  
DB 191 LGLTGVALLQLLLPFCPSPRYLLINKNEARAKKALQRLGTADVDSQEVAKMDES 250  
QY 295 VQRIRVSVLELLRAPPYVQVVTIVTMACYQLCGNNAIFVYTNISIFGKAGIPPAKIP 354  
DB 251 XMXSEKVSULELFRSRXYROPVIAIVLQLSQOLSGINAVFYVYSTSIFEKAGV--QGPV 308  
QY 355 YVTLSTGGIEITLAAVFSGLVIEHLGRPLIGGFLMGLPFGTLTITLTLQDHAPWVP 414  
DB 309 YATIGAGVNVTVFVSVFVVERAGRRLTLHLGLGGMAGCAVLMTIALLLDQVPMWSYV 368  
QY 415 SIVGILAIASFCSPGGIPILTGEFFQOSQORPAARIIAGTVNWLNSNEAVGLLFPPIQK 474  
DB 369 SIVAIFGVAFVFGPPIPMFIVAEILFSQGPRAATAVAGFSNWTNFIIVGLLFOVIAE 428  
QY 475 SLDTYCYFLVFATCITGAIYLYFVLPETKRTYAEISQAFSKRKNKAYPPEEK 526  
DB 429 LLGPVVFVFAVALLLFFITFLVAVPDKGTPTDEIAAFAFKXKXQPEKE 480

RESULT 8

US-09-299-549-10  
; Sequence 10, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; APPLICANT: Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; TITLE OF INVENTION: GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 10:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 493 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-10

Query Match 37.2%; Score 1081; DB 3; Length 493;  
Best Local Similarity 48.5%; Pred. No. 7.8e-99;  
Matches 229; Conservative 71; Mismatches 168; Indels 4; Gaps 3;

Db 433 QEGLGPSFVFAVICLITTIYIFLVIPETKAKTFIEINQIFTKMNKSEVYPEKEEL 490  
Db 317 VTAGTGVNVMVTFCAVF---VVELLGRLLLLGFCICLIACCVLTAALALQDTVSWMP 373  
QY 413 YLSIVGILAIATSCSGPGGPFILTBGEFFQOSORPAAFIAGTVNWSNFAVGLLPFFI 472  
Db 374 YISIVCVISYVIGHALGPSIPALLI-TIFQSRPSAFWVGSVHWSNFTVGLLPFFI 432  
QY 473 QKSLDTYCYFLVFATICITGAIYLYFLVPETKNTYAEISQAFSKRNK---AYPEEKI 527  
Db 433 QEGLGPSFVFAVICLITTIYIFLVIPETKAKTFIEINQIFTKMNKSEVYPEKEEL 490

Db 433 QEGLGPSFVFAVICLITTIYIFLVIPETKAKTFIEINQIFTKMNKSEVYPEKEEL 490  
Db 317 VTAGTGVNVMVTFCAVF---VVELLGRLLLLGFCICLIACCVLTAALALQDTVSWMP 373  
QY 413 YLSIVGILAIATSCSGPGGPFILTBGEFFQOSORPAAFIAGTVNWSNFAVGLLPFFI 472  
Db 374 YISIVCVISYVIGHALGPSIPALLI-TIFQSRPSAFWVGSVHWSNFTVGLLPFFI 432  
QY 473 QKSLDTYCYFLVFATICITGAIYLYFLVPETKNTYAEISQAFSKRNK---AYPEEKI 527  
Db 433 QEGLGPSFVFAVICLITTIYIFLVIPETKAKTFIEINQIFTKMNKSEVYPEKEEL 490

RESULT 6  
US-09-610-417-7  
; Sequence 7, Application US/09610417  
; Patent No. 6346374  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/610,417  
; FILING DATE: 05-Jul-2000  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/299,549  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 7:  
US-09-610-417-7

Query Match 38.2%; Score 1108; DB 3; Length 500;  
Best Local Similarity 46.2%; Pred. No. 1.6e-101;  
Matches 221; Conservative 92; Mismatches 155; Indels 10; Gaps 4;  
QY 56 LLVASLAGAGSSFLYGNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLMSVTVS 115  
Db 17 LALATLIAAFGSSFOYGVNVAVNSPALLMQQFYNETYTGRTGFEFMDPFLTLMSVTVS 76  
QY 116 IFAIGGLVGTLLVKNMGVLGRKHTLLANNQFAISALLMACSLQAGAFEMLVGRIMG 175  
Db 77 MFPPFGFTGSLVGLVPLNKNKGRKALLFNFIIVPALMGCSRVATSFELIIISRLVG 136  
QY 176 IDGGVALSVLPMYLSISPKEIRGSLGQVTAIFTCIGVFTGQLLGLPELLGKESWPYLF 235  
Db 137 ICAGVSNVPMYLGELAPKVLGALGVWPQLFTVIGILVAQIFGLNLLANVDGWPILL 196  
QY 236 GVIIVPAVQLSLPFLPDSPPRYLLLEKNEARAVKAFQTPFLGKADYSQVEEVLAEHV 295  
Db 197 GLTGVPALQLLPLFPFESPRYLLIOKDBAAAKALQTLRGWDSVDREVAERQDEA 256  
QY 296 QRSRLVSVLELLKAPYVRQVTVIVTMACYQLCGLNATWFTNYSIFGAGIPPAKIPY 355  
Db 257 EKAAGFISVLKFRMSLRWQLLSIIVLMGGQQLSGVNAIYYADQIYLSAGVPEEHVQY 316  
QY 356 VTLSTGGIE---TLAAVPSGLVIEHLGRRPLLIGGFGMLGFFCTLTITLTLDQADWPV 412

Query Match 38.2%; Score 1108; DB 4; Length 500;  
Best Local Similarity 46.2%; Pred. No. 1.6e-101;  
Matches 221; Conservative 92; Mismatches 155; Indels 10; Gaps 4;  
QY 56 LLVASLAGAGSSFLYGNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLMSVTVS 115  
Db 17 LALATLIAAFGSSFOYGVNVAVNSPALLMQQFYNETYTGRTGFEFMDPFLTLMSVTVS 76  
QY 116 IFAIGGLVGTLLVKNMGVLGRKHTLLANNQFAISALLMACSLQAGAFEMLVGRIMG 175  
Db 77 MFPPFGFTGSLVGLVPLNKNKGRKALLFNFIIVPALMGCSRVATSFELIIISRLVG 136  
QY 176 IDGGVALSVLPMYLSISPKEIRGSLGQVTAIFTCIGVFTGQLLGLPELLGKESWPYLF 235  
Db 137 ICAGVSNVPMYLGELAPKVLGALGVWPQLFTVIGILVAQIFGLNLLANVDGWPILL 196

Query Match 38.2%; Score 1108; DB 3; Length 500;  
Best Local Similarity 46.2%; Pred. No. 1.6e-101;  
Matches 221; Conservative 92; Mismatches 155; Indels 10; Gaps 4;  
QY 56 LLVASLAGAGSSFLYGNLSVNVNAPTPIKAFYNESWERRHGRPIDPDTLLMSVTVS 115  
Db 17 LALATLIAAFGSSFOYGVNVAVNSPALLMQQFYNETYTGRTGFEFMDPFLTLMSVTVS 76  
QY 116 IFAIGGLVGTLLVKNMGVLGRKHTLLANNQFAISALLMACSLQAGAFEMLVGRIMG 175  
Db 77 MFPPFGFTGSLVGLVPLNKNKGRKALLFNFIIVPALMGCSRVATSFELIIISRLVG 136  
QY 176 IDGGVALSVLPMYLSISPKEIRGSLGQVTAIFTCIGVFTGQLLGLPELLGKESWPYLF 235  
Db 137 ICAGVSNVPMYLGELAPKVLGALGVWPQLFTVIGILVAQIFGLNLLANVDGWPILL 196  
QY 236 GVIIVPAVQLSLPFLPDSPPRYLLLEKNEARAVKAFQTPFLGKADYSQVEEVLAEHV 295  
Db 197 GLTGVPALQLLPLFPFESPRYLLIOKDBAAAKALQTLRGWDSVDREVAERQDEA 256  
QY 296 QRSRLVSVLELLKAPYVRQVTVIVTMACYQLCGLNATWFTNYSIFGAGIPPAKIPY 355  
Db 257 EKAAGFISVLKFRMSLRWQLLSIIVLMGGQQLSGVNAIYYADQIYLSAGVPEEHVQY 316  
QY 356 VTLSTGGIE---TLAAVPSGLVIEHLGRRPLLIGGFGMLGFFCTLTITLTLDQADWPV 412

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-299-549-7  
; Sequence 7, Application US/09299549  
; Patent No. 6136547  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/299,549  
; FILING DATE: 26-APR-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 500 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-299-549-7

US-09-031-392-7  
 ; Sequence 7, Application US/09031392  
 ; Patent No. 5942398  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tartaglia, Louis A.  
 ; APPLICANT: Weng, Xun  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES  
 ; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF  
 ; NUMBER OF SEQUENCES: 10  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Fish & Richardson P.C.  
 ; STREET: 225 Franklin Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110-2804  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: Windows95  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/031,392  
 ; FILING DATE: 26-FEB-1998  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Meiklejohn, Ph.D., Anita L.  
 ; REGISTRATION NUMBER: 35,283  
 ; REFERENCE/DOCKET NUMBER: 07334/072001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617/542-5070  
 ; TELEFAX: 617/542-8906  
 ; TELEX: 200154  
 ; INFORMATION FOR SEQ ID NO: 7:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 500 amino acids  
 ; TYPE: amino acid  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; US-09-031-392-7

Query Match 38.2%; Score 1108; DB 2; Length 500;  
 Best Local Similarity 46.2%; Pred. No. 1.6e-101;  
 Matches 221; Conservative 92; Mismatches 155; Indels 10; Gaps 4;

QY	56	LLVASLAGAFSSFLYGNLSVNAAPTYIKAFNYESWERRHGRPIDDTLTLLWSVTVS	115
Db	17	LALATLIAAFGSSFOYGYNVAANVSPALLMQOFTYGTGTGTFDFDTLLWSVTVS	76
QY	116	IFATGLVGLTVKVMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFTMG	175
Db	77	MFPFGFTGSLVGLVPLVKNFKRGKALLFNNTFIVPAILMCCSRVATSFELIISRLIVG	136
QY	176	IDGGVALSVLPMYLSLSEISPKIRSGVGTAFICIGVFTGQLLGLPELLGKSTWPLYF	235
Db	137	ICAGVSSNVPMYLGELAPKNLRGALGVVPOLFITVGLVVAIFGLRNLLANVDGWPILL	196
QY	236	GVIVPVAVQVLLSPLPDPSPRYLLLEKHNARAVKAFQTFGLKADVSQVEEVLAESHV	295
Db	197	GLTGVPALQVLLLPFPFESPRYLLIQKDEAAKALQTLRGWDSVDREVAIRQDEA	256
QY	296	QRSIRLVSVLELLRAPYVRQWVTVTMACYQLGLNAINWFTYNSIFGKAGIPAKIPY	355
Db	257	EKAAGFISVLYKLFMRSLRWQLLSIIVLMGGQSLGVNAIYYADQIYLSAGVPEEHVQY	316
QY	356	VTLSGGE---TLAAVPSGLVIEHGERPILLIGFGLMGLFFGTTLTTLQDHAPWVP	412
Db	317	VTAGTGVNVVMTFCAVF---VVELLGRRLLLGLGFSICLIACCVLTAALALQDTSWMP	373
QY	413	YLSIVGILAIASFCSPGGIPFILTGFQQSORPAAFIAGTVNMLSNFVAGLLPFPFI	472
Db	374	YISIVCVISVIGHALGSPDALLI-TIFQSSPSPAFMVGSGVHMLSNVTGLIFPFI	432
QY	473	QKSLDTYCFVLPATICITGAIYLYFVLPETKNTVAEISQAFSKRNK---AYPPEKI	527

COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows95  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/610,417  
 FILING DATE: 05-Jul-2000  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 09/299,549  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Meiklejohn, Ph.D., Anita L.  
 REGISTRATION NUMBER: 35,283  
 REFERENCE/DOCKET NUMBER: 07334/072002  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617/542-5070  
 TELEFAX: 617/542-8906  
 TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 563 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FRAGMENT TYPE: internal  
 SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-610-417-2

Query Match 100.0%; Score 2903; DB 4; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-280;  
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MARKQNRNSELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRKQWSCSILLVAS	60
Db	1	MARKQNRNSELGLVPLTDDTSHAGPPGGRALLECDHLRSVPGGRRKQWSCSILLVAS	60
QY	61	LACAFGSSFLYGNLSVNAAPTYIKAFNYESWERRHGRPIDDTLTLLWSVTVSFAIG	120
Db	61	LACAFGSSFLYGNLSVNAAPTYIKAFNYESWERRHGRPIDDTLTLLWSVTVSFAIG	120
QY	121	GLVGLTVKVMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFTMGIDGV	180
Db	121	GLVGLTVKVMIGKVLGRKHTLLANNFAISALLMACSLQAGAFEMLIYGRFTMGIDGV	180
QY	181	ALSVLPMYLSLSEISPKIRSGVGTAFICIGVFTGQLLGLPELLGKSTWPLYFGVIVV	240
Db	181	ALSVLPMYLSLSEISPKIRSGVGTAFICIGVFTGQLLGLPELLGKSTWPLYFGVIVV	240
QY	241	PAVVQLLSPLPDPSPRYLLLEKHNARAVKAFQTFGLKADVSQVEEVLAESHVQSR	300
Db	241	PAVVQLLSPLPDPSPRYLLLEKHNARAVKAFQTFGLKADVSQVEEVLAESHVQSR	300
QY	301	LVSVLELLRAPYVRQWVTVTMACYQLGLNAINWFTYNSIFGKAGIPAKIPYVTLST	360
Db	301	LVSVLELLRAPYVRQWVTVTMACYQLGLNAINWFTYNSIFGKAGIPAKIPYVTLST	360
QY	361	GGIETLAAVPSGLVIEHGERPILLIGFGLMGLFFGTTLTTLQDHAPWVPYLSIVGIL	420
Db	361	GGIETLAAVPSGLVIEHGERPILLIGFGLMGLFFGTTLTTLQDHAPWVPYLSIVGIL	420
QY	421	AIIASFCSGGGIPFILTGTFFQSORPAAFIAGTVNMLSNFVAGLLPFPFIQKSLDTYC	480
Db	421	AIIASFCSGGGIPFILTGTFFQSORPAAFIAGTVNMLSNFVAGLLPFPFIQKSLDTYC	480
QY	481	FLVPATICITGAIYLYFVLPETKNTVAEISQAFSKRNKAYPPEEKIDSAVTPASSPF	540
Db	481	FLVPATICITGAIYLYFVLPETKNTVAEISQAFSKRNKAYPPEEKIDSAVTPASSPF	540
QY	541	TPNTAWIQAAATTTATKKEHPL	563
Db	541	TPNTAWIQAAATTTATKKEHPL	563

QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLLTLLWSVTVSIFAIG 120  
DB 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLLTLLWSVTVSIFAIG 120  
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180  
DB 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180  
QY 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWVYLFVIV 240  
DB 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWVYLFVIV 240  
QY 241 PAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEHVSIR 300  
DB 241 PAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEHVSIR 300  
QY 301 LVSVELELRAPYVRQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360  
DB 301 LVSVELELRAPYVRQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360  
QY 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLLTTLQDHAPWVYLSIVGIL 420  
DB 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLLTTLQDHAPWVYLSIVGIL 420  
QY 421 AIIASFCSGPGGIPFLTGBFFQSQRPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 480  
DB 421 AIIASFCSGPGGIPFLTGBFFQSQRPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 480  
QY 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540  
DB 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540  
QY 541 TTPNTAWIOAAATTTATKKEHPL 563  
DB 541 TTPNTAWIOAAATTTATKKEHPL 563

RESULT 2

US-09-299-549-2  
; Sequence 2, Application US/09299549

; Patent No. 6136547

; GENERAL INFORMATION:

; APPLICANT: Tartaglia, Louis A.

; APPLICANT: Weng, Xun

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING

; TITLE OF INVENTION: GLUTEX AND USES THEREOF

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: Windows95

; SOFTWARE: FASTSEQ for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/299,549

; FILING DATE: 26-APR-1999

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 09/031,392

; FILING DATE: 26-FEB-1998

; ATTORNEY/AGENT INFORMATION:

; NAME: Meiklejohn, Ph.D., Anita L.

; REGISTRATION NUMBER: 35,283

; REFERENCE/DOCKET NUMBER: 07334/072002

; TELEPHONE: 617/542-5070

; TELEFAX: 617/542-8906

; TELEX: 200154

INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 563 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; FRAGMENT TYPE: internal

US-09-299-549-2

Query Match 100.0%; Score 2903; DB 3; Length 563;

Best Local Similarity 100.0%; Pred. No. 3.9e-280;

Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGPGRALLLECHLRSQVPGRRRKDWSCSLIVAS 60

DB 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGPGRALLLECHLRSQVPGRRRKDWSCSLIVAS 60

QY 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLLTLLWSVTVSIFAIG 120

DB 61 LAGAFSSFLYGNLSVNNAPTPIKAFYNESWERRHGRPIDDTLLTLLWSVTVSIFAIG 120

QY 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180

DB 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAISAAALLMACSLQAGAFEMLIIVGRFIMGIDGV 180

QY 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWVYLFVIV 240

DB 181 ALSVLPWYLSEISPKIRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWVYLFVIV 240

QY 241 PAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEHVSIR 300

DB 241 PAVQVLLSLPFLPDSPRYLLEKHNARAVKAFQTLFGKADVSQVEEVLAEHVSIR 300

QY 301 LVSVELELRAPYVRQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360

DB 301 LVSVELELRAPYVRQVTVITMACYQLCGLNAINWFTNSIFGKAGIPPAKIPYVTLST 360

QY 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLLTTLQDHAPWVYLSIVGIL 420

DB 361 GGIETLAAVFSGLVIEHLGRRLPILGFGMLGFFGTLLTTLQDHAPWVYLSIVGIL 420

QY 421 AIIASFCSGPGGIPFLTGBFFQSQRPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 480

DB 421 AIIASFCSGPGGIPFLTGBFFQSQRPAAFIAGTVNWSNFAVGLLFPFIQKSLDTC 480

QY 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540

DB 481 FLVFATCITGAIYLYFVLPETKNTYAEISQAFSKENKAYPPEEKIDSAVTDAPASSPF 540

QY 541 TTPNTAWIOAAATTTATKKEHPL 563

DB 541 TTPNTAWIOAAATTTATKKEHPL 563

RESULT 3

US-09-610-417-2

; Sequence 2, Application US/09610417

; Patent No. 6346374

; GENERAL INFORMATION:

; APPLICANT: Tartaglia, Louis A.

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING

; GLUTEX AND USES THEREOF

; NUMBER OF SEQUENCES: 10

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fish & Richardson P.C.

; STREET: 225 Franklin Street

; CITY: Boston

; STATE: MA

; COUNTRY: USA

; ZIP: 02110-2804

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

OM protein - protein search, using sw model

Run on: April 6, 2004, 10:15:52 ; Search time 23 seconds  
(without alignments)  
1263.714 Million cell updates/sec

Title: US-09-981-947B-2  
Perfect score: 2903  
Sequence: 1 MARKQNSKELGLVPLTDD.....NTAWIQAAATTATKKEHPL 563

Scoring table:  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgm2\_6/ptodata/2/iaa/5A.COMB.pep:\*  
2: /cgm2\_6/ptodata/2/iaa/5B.COMB.pep:\*  
3: /cgm2\_6/ptodata/2/iaa/6A.COMB.pep:\*  
4: /cgm2\_6/ptodata/2/iaa/6B.COMB.pep:\*  
5: /cgm2\_6/ptodata/2/iaa/PTUS.COMB.pep:\*  
6: /cgm2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2903	100.0	563	2	US-09-031-392-2
2	2903	100.0	563	3	US-09-299-549-2
3	2903	100.0	563	4	US-09-610-417-2
4	1108	38.2	500	2	US-09-031-392-7
5	1108	38.2	500	3	US-09-299-549-7
6	1108	38.2	500	4	US-09-610-417-7
7	1081	37.2	493	2	US-09-031-392-10
8	1081	37.2	493	3	US-09-299-549-10
9	1081	37.2	493	4	US-09-610-417-10
10	919.5	31.7	492	2	US-08-355-844-3
11	919.5	31.7	492	5	PT-US95-16126-3
12	874	30.1	509	2	US-09-031-392-6
13	874	30.1	509	3	US-09-299-549-6
14	874	30.1	509	4	US-09-610-417-6
15	866	29.8	494	2	US-09-031-392-5
16	866	29.8	494	3	US-09-299-549-5
17	866	29.8	494	4	US-09-610-417-5
18	823.5	28.4	863	4	US-09-894-927B-9
19	751	25.9	534	2	US-09-031-392-4
20	751	25.9	534	3	US-09-299-549-4
21	751	25.9	534	4	US-09-610-417-4
22	725.5	25.0	383	2	US-09-031-392-3
23	725.5	25.0	383	3	US-09-299-549-3
24	725.5	25.0	383	4	US-09-610-417-3
25	695.5	24.0	524	2	US-08-928-692-12
26	695.5	24.0	524	4	US-09-339-972-12
27	600	20.7	488	4	US-10-162-012-46

Sequence 11, Appl  
Sequence 11, Appl  
Sequence 30, Appl  
Sequence 24, Appl  
Sequence 11933, A  
Sequence 11731, A  
Sequence 22, Appl  
Sequence 2, Appl  
Sequence 28, Appl  
Sequence 19, Appl  
Sequence 11902, A  
Sequence 10, Appl  
Sequence 10, Appl  
Sequence 22, Appl  
Sequence 13, Appl  
Sequence 13, Appl  
Sequence 26, Appl  
Sequence 23, Appl

28 480 16.5 488 2 US-08-928-692-11  
29 480 16.5 488 4 US-09-339-972-11  
30 422.5 14.6 549 4 US-09-291-922-30  
31 419 14.4 523 4 US-09-291-922-24  
32 418.5 14.4 476 4 US-09-489-039A-11933  
33 417.5 14.4 501 4 US-09-489-039A-11731  
34 397.5 13.7 510 4 US-09-291-922-22  
35 396.5 13.7 502 4 US-09-679-686B-2  
36 390 13.4 529 4 US-09-291-922-28  
37 389.5 13.4 510 4 US-09-679-686B-19  
38 389 13.4 514 4 US-09-489-039A-11902  
39 382.5 13.2 488 2 US-08-928-692-10  
40 382.5 13.2 488 4 US-09-339-972-10  
41 382 13.2 514 4 US-09-679-686B-22  
42 374.5 12.9 584 2 US-08-928-692-13  
43 374.5 12.9 584 4 US-09-339-972-13  
44 372 12.8 539 4 US-09-291-922-26  
45 369 12.7 518 4 US-09-679-686B-23

ALIGNMENTS

RESULT 1  
US-09-031-392-2  
; Sequence 2, Application US/09031392  
; Patent No. 5942398  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; APPLICANT: Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/031,392  
; FILING DATE: 26-FEB-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Meiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 563 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-09-031-392-2

Query Match 100.0%; Score 2903; DB 2; Length 563;  
Best Local Similarity 100.0%; Pred. No. 3.9e-280;  
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MARKQNSKELGLVPLTDDTSHAGPPGCRALLECHLRSVPGGRRRDKWSCSLIVAS 60  
Db 1 MARKQNSKELGLVPLTDDTSHAGPPGCRALLECHLRSVPGGRRRDKWSCSLIVAS 60





Best Local Similarity 99.5%; Pred. No. 3.3e-196; Mismatches 2; Indels 0; Gaps 0; Matches 395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 137 RKHTLLANNNGFPAISAAALLMACSLQAGAFEMLVIVGRFIMGIDGGVALSVLPMLSEISPK 196

Db 5 RKHTLLANNNGFPAISAAALLMACSLQAGAFEMLVIVGRFIMGIDGGVALSVLPMLSEISPK 64

QY 197 IRGSLGQVTAIFTCIGVFTGQLLGLPELIGKSTWPLYFGVIVPAVQVLLSLPLPDPSP 256

Db 65 IRGSLGQVTAIFTCIGVFTGQLLGLPELIGKSTWPLYFGVIVPAVQVLLSLPLPDPSP 124

QY 257 RYLLLEKHEARAVKAPQTFGLKADVSQVEVEVLAESHVQSRIRLVLELLRAPVVRWQ 316

Db 125 RYLLLEKHEARAVKAPQTFGLKADVSQVEVEVLAESHVQSRIRLVLELLRAPVVRWQ 184

QY 317 VVTVIVTMACYOICGLNAINFYTNSIFGKAGIPPAKIPVVTLSGTGIELTAAVFSGLVIE 376

Db 185 VVTVIVTMACYOICGLNAINFYTNSIFGKAGIPPAKIPVVTLSGTGIELTAAVFSGLVIE 244

QY 377 HLGRRPLLIGFGGLMGFFGTITLTLLQDHAPWPVLSIVGILAIASFCSGGGIPFI 436

Db 245 HLGRRPLLIGFGGLMGFFGTITLTLLQDHAPWPVLSIVGILAIASFCSGGGIPFI 304

QY 437 LTGFEFFQSQRPAAFIAGTVNMLSNFAVGLLPFFIOKSLDTYCFVLFATICTGAIYLY 496

Db 305 LTGFEFFQSQRPAAFIAGTVNMLSNFAVGLLPFFIOKSLDTYCFVLFATICTGAIYLY 364

QY 497 FVLPEKNTVABISQAFSKRNKAYPPEEKIDSAVTD 533

Db 365 FVLPEKNTVABISQAFSKRNKAYPPEEKIDSAVTD 401

RESULT 15

AAE04893

ID AAE04893 standard; protein; 416 AA.

XX AC AAE04893;

XX DT 10-SEP-2001 (first entry)

XX DE Human transporter and ion channel-6 (TRICH-6) protein.

XX KW Human; transporter and ion channel-6; TRICH-6; vaccine; cystic fibrosis; gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy; hypertension; angina; neurological disorder; asthma; bipolar disorder; dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia; Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke; Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy; demyelinating disease; mental disorder; Schizophrenia; polyomyelitis; muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease; dermatomyositis; diabetes mellitus; immunological disorder; psoriasis; rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus; sickle cell anaemia; Wilson's disease; infertility; Cushing's disease; scleroderma; pulmonary artery stenosis; neutropenic; Addison's disease; malabsorption syndrome; hypercholesterolaemia; cancer.

XX OS Homo sapiens.

XX FH Key

FT Peptide 1..37

FT Domain /label= Signal\_peptide 30..416

FT Protein /note= "Sugar transporter domain"

FT Domain /label= Mature\_human\_TRICH-6\_protein 38..416

FT Domain /label= Transmembrane\_domain 83..99

FT Domain /label= Transmembrane\_domain 356..375

XX WO200146258-A2.

XX PD 28-JUN-2001.

QY 497 FVLPEKNTVABISQAFSKRNKAYPPEEKIDSAVTD 533

Db 371 FVLPEKNTVABISQAFSKRNKAYPPEEKIDSAVTD 407

RESULT 14

AAM39530

ID AAM39530 standard; protein; 408 AA.

XX AC AAM39530;

XX DT 22-OCT-2001 (first entry)

XX DE Human polypeptide SEQ ID NO 2675.

XX KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer; peripheral nervous system; neuropathy; central nervous system; CNS; Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic; amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic; chemokinetic; thrombolytic; drug screening; arthritis; inflammation; leukaemia.

XX OS Homo sapiens.

XX PN WO200153312-A1.

XX PD 26-JUL-2001.

XX PF 26-DEC-2000; 2000WO-US034263.

XX PR 23-DEC-1999; 99US-00471275.

XX PR 21-JAN-2000; 2000US-00488725.

XX PR 25-APR-2000; 2000US-00552317.

XX PR 20-JUN-2000; 2000US-00598042.

XX PR 19-JUL-2000; 2000US-00620312.

XX PR 03-AUG-2000; 2000US-00653450.

XX PR 14-SEP-2000; 2000US-00662191.

XX PR 19-OCT-2000; 2000US-00693036.

XX PR 29-NOV-2000; 2000US-00727344.

XX PA (HYSE-) HYSEQ INC.

XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D; Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA; Zhou P, Goodrich R, Drmanac RT;

XX WPI; 2001-442253/47.

XX N-PSDB; AAI59686.

XX Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries.

XX Example 4; SEQ ID NO 2675; 10078pp; English.

XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the encoded polypeptides (AAM38642-AAM42213) with nootropic, immunosuppressant and cytostatic activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and localised neuropathies and central nervous system diseases, such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. Other uses include the utilisation of the activities such as: immune system suppression, activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukaemias and C.N.S disorders. Note: The sequence data for this patent did not form part of the printed specification

XX Sequence 408 AA;

Query Match 69.3%; Score 2012; DB 4; Length 408;

XX SQ Sequence 408 AA; Query Match 69.5%; Score 2018; DB 4; Length 408; Best Local Similarity 99.5%; Pred. No. 8e-197; Mismatches 1; Indels 0; Gaps 0; Matches 395; Conservative

QY 137 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 196  
DB 5 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 64  
QY 197 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWYLFVGVIVPAVVQVLLSLFPDPSF 256  
DB 65 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWYLFVGVIVPAVVQVLLSLFPDPSF 124  
QY 257 RYLLEKHNREARAVKAFQTFGLKADVSQVEEVLAEHVORSIRLVSVLELLRAPVVRWQ 316  
DB 125 RYLLEKHNREARAVKAFQTFGLKADVSQVEEVLAEHVORSIRLVSVLELLRAPVVRWQ 184  
QY 317 VVTIVTMACYQLCGLNAINFWYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 376  
DB 185 VVTIVTMACYQLCGLNAINFWYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 244  
QY 377 HLGRRPILLIGGFGMLGFFGTLTITLTLODHAPWPVYLSIVGILAIIASFCSPGGGIPPI 436  
DB 245 HLGRRPILLIGGFGMLGFFGTLTITLTLODHAPWPVYLSIVGILAIIASFCSPGGGIPPI 304  
QY 437 LTGEFFQOSORPAAFIAGTVNWSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 496  
DB 305 LTGEFFQOSORPAAFIAGTVNWSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 364  
QY 497 FVLPEKTRNRYVAEISQAFSKENKAYPPEEKIDSAVTD 533  
DB 365 FVLPEKTRNRYVAEISQAFSKENKAYPPEEKIDSAVTD 401

RESULT 13  
IDC AD26265 standard; protein; 422 AA.  
XX AC AD26265;  
XX AC  
XX DT 18-DEC-2003 (first entry)  
XX DE Human NOV17h protein.  
XX KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;  
XX KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;  
XX KW transgenic; human.  
XX OS Homo sapiens.  
XX OS  
XX EN WO2003004687-A2.  
XX PD 16-JAN-2003.  
XX PF  
XX PF 03-JUL-2002; 2002WO-US021361.  
XX PR 05-JUL-2001; 2001US-0303046P.  
XX PR 09-JUL-2001; 2001US-0303828P.  
XX PR 09-JUL-2001; 2001US-0304016P.  
XX PR 11-JUL-2001; 2001US-0304502P.  
XX PR 13-JUL-2001; 2001US-0305262P.  
XX PR 16-JUL-2001; 2001US-0305673P.  
XX PR 17-JUL-2001; 2001US-0306085P.  
XX PR 24-JUL-2001; 2001US-0307536P.  
XX PR 27-JUL-2001; 2001US-0308228P.  
XX PR 30-JUL-2001; 2001US-0308877P.  
XX PR 01-AUG-2001; 2001US-0309255P.  
XX PR 17-AUG-2001; 2001US-0313328P.  
XX PR 12-SEP-2001; 2001US-0318711P.  
XX PR 19-SEP-2001; 2001US-0323380P.  
XX PR 21-SEP-2001; 2001US-0323969P.

PR 04-JAN-2002; 2002US-0345022P.  
PR 04-JAN-2002; 2002US-0345038P.  
PR 28-FEB-2002; 2002US-0361172P.  
PR 01-MAR-2002; 2002US-0360814P.  
PR 01-MAR-2002; 2002US-0360830P.  
PR 01-MAR-2002; 2002US-0361133P.  
PR 01-MAR-2002; 2002US-0361147P.  
PR 05-MAR-2002; 2002US-0361677P.  
PR 12-APR-2002; 2002US-0363637P.  
PR 16-APR-2002; 2002US-0372990P.  
PR 19-APR-2002; 2002US-0373881P.  
PR 19-APR-2002; 2002US-0373921P.  
PR 02-JUL-2002; 2002US-00188186.  
XX (CURA-) CUBAGEN CORP.  
XX PA Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ, Catterton B, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Gorman L; Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE; Padigaru M, Patturajan M, Pena CE, Rastelli L, Shenoy S; Shimketa R, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ; Vernet CAM, Voss EZ, Zhong M; WPI; 2003-221607/21.  
DR N-PSDB; AD26264.  
XX New isolated NOVX polypeptide, useful for determining the presence of, or predisposition to a disease associated with altered levels of expression of the polypeptide, and for treating or preventing cancer.  
XX Claim 2; SEQ ID NO 90; 478pp; English.  
XX The invention relates to a novel isolated NOV polypeptide. The polypeptide of the invention demonstrates cytosolic activity and may be used for determining the presence of, or predisposition to a disease associated with altered levels of expression of the polypeptide, including metabolic disorders, immune disorders, neurodegenerative disorders, circulatory diseases, haemopoietic disorders, wasting diseases and cancer. The polypeptide may also be utilised during gene therapy procedures, vaccine development and transgenic animal production. The current sequence is that of the human NOV protein of the invention.

XX SQ Sequence 422 AA; Query Match 69.5%; Score 2017; DB 7; Length 422; Best Local Similarity 99.5%; Pred. No. 1.1e-196; Matches 395; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 137 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 196  
DB 11 RKHTLLANNNGFAISAALLMACSLQAGAFEMLIIVGRFTMGIDGGVALSVLPYLSISPKE 70  
QY 197 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWYLFVGVIVPAVVQVLLSLFPDPSF 256  
DB 71 IRGSLGQVTAIFICIGVFTGQLLGLPELLGKSTWYLFVGVIVPAVVQVLLSLFPDPSF 130  
QY 257 RYLLEKHNREARAVKAFQTFGLKADVSQVEEVLAEHVORSIRLVSVLELLRAPVVRWQ 316  
DB 131 RYLLEKHNREARAVKAFQTFGLKADVSQVEEVLAEHVORSIRLVSVLELLRAPVVRWQ 190  
QY 317 VVTIVTMACYQLCGLNAINFWYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 376  
DB 191 VVTIVTMACYQLCGLNAINFWYNSIPGKAGIPPAKIPYVTLSTGGIETLAIVFSGLVIE 250  
QY 377 HLGRRPILLIGGFGMLGFFGTLTITLTLODHAPWPVYLSIVGILAIIASFCSPGGGIPPI 436  
DB 251 HLGRRPILLIGGFGMLGFFGTLTITLTLODHAPWPVYLSIVGILAIIASFCSPGGGIPPI 310  
QY 437 LTGEFFQOSORPAAFIAGTVNWSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 496  
DB 311 LTGEFFQOSORPAAFIAGTVNWSNFAVGLLPFIQKSLDTCFLVFAICITGAIYLY 370

PR 05-MAR-2002; 2002US-0361677P.  
 PR 02-APR-2002; 2002US-0363637P.  
 PR 12-APR-2002; 2002US-0372326P.  
 PR 16-APR-2002; 2002US-0372990P.  
 PR 19-APR-2002; 2002US-0373881P.  
 PR 19-APR-2002; 2002US-0373921P.  
 PR 02-JUL-2002; 2002US-00198186.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Anderson DW, Berghs C, Boidog FL, Burgess CE, Casman SJ, Gorman L;  
 PI Catterson E, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Miller CE;  
 PI Guo X, Jeffers M, Kekuda R, Li L, Malvankar UM, Miller CE;  
 PI Radigaru M, Patturajan M, Pena CE, Rastelli L, Shenoy S;  
 PI Shimkets RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ;  
 PI Vernet CAM, Voss EZ, Zhong M;  
 XX  
 XX WPI: 2003-221607/21.  
 DR N-PSDB; ADC26252.  
 XX  
 XX New isolated NOVX polypeptide, useful for determining the presence of, or  
 PT predisposition to a disease associated with altered levels of expression  
 PT of the polypeptide, and for treating or preventing cancer.  
 XX  
 XX Claim 2; SEQ ID NO 78; 478pp; English.  
 XX  
 XX The invention relates to a novel isolated NOV polypeptide. The  
 CC polypeptide of the invention demonstrates cytostatic activity and may be  
 CC used for determining the presence of, or predisposition to a disease  
 CC associated with altered levels of expression of the polypeptide,  
 CC including metabolic disorders, immune disorders, neurodegenerative  
 CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases  
 CC and cancer. The polypeptide may also be utilised during gene therapy  
 CC procedures, vaccine development and transgenic animal production. The  
 CC current sequence is that of the human NOV protein of the invention.  
 XX  
 XX Sequence 565 AA;  
 XX  
 XX Query Match 80.5%; Score 2337.5; DB 7; Length 565;  
 XX Best Local Similarity 77.3%; Pred. No. 3e-229;  
 XX Matches 474; Conservative 0; Mismatches 4; Indels 135; Gaps 2;  
 QY 1 MARKQNSKELGLVPLTDDTSHAGPPGPGRALLECDHLRSYGVGGRRK----- 50  
 DB 1 MARKQNSKELGLVPLTDDTSHAGPPGPGRALLECDHLRSYGVGGRRKPLRSTSSAA 60  
 QY 51 ----- 50  
 DB 61 GSSTTYVASAAIKPIPHRWASLAPEKSHLLRLQLRGLALEVKLISLGSKPOSSISWL 120  
 QY 51 -----DWSCLLVASLAGAFSGSFLYGNLSVYNAPTPIYKAFYNESWERRGR 100  
 DB 121 ALEETNEIGRDWSCLLVASLAGAFSGSFLYGNLSVYNAPT----- 163  
 QY 101 IDPPTLLMSVTVSIFAIGLVOTLIVKMGKVLGRKHLLANNNGPAISAAALLMACSLQ 160  
 DB 164 -----HTLLANNNGPAISAAALLMACSLQ 185  
 QY 161 AGAFEMLIIVGRFIMGIDGGVALSVLPMVLSEISPEKIRSGVQTAFICIGVFTGQLG 220  
 DB 186 AGAFEMLIIVGRFIMGIDGGVALSVLPMVLSEISPEKIRSGVQTAFICIGVFTGQLG 245  
 QY 221 LPELLKGESTWPLYFGVIVPAVQVLLSLPLPDSPRYLLLEKNEARAVKAFQTFLGKA 280  
 DB 246 LPELLKGESTWPLYFGVIVPAVQVLLSLPLPDSPRYLLLEKNEARAVKAFQTFLGKA 305  
 QY 281 DVSQVEVEVLAESHVQSRIRLVSVLELLRAPYVRQVVTIVTVACVQLCGLNIAWFTYN 340  
 DB 306 DVSQVEVEVLAESHVQSRIRLVSVLELLRAPYVRQVVTIVTVACVQLCGLNIAWFTYN 365  
 QY 341 SIFGKAGIPPAKIPVYTLSTGGIETLAAVFSGLVIEHGRPLLIIGFGLMGLFPFGLTI 400  
 DB 366 SIFGKAGIPPAKIPVYTLSTGGIETLAAVFSGLVIEHGRPLLIIGFGLMGLFPFGLTI 425

QY 401 TLTLDHAPWVPLSVIGILAIITASFCSGPGGIPFILTGEFFQOSORPAAFIAGTVNWL 460  
 DB 426 TLTLDHAPWVPLSVIGILAIITASFCSGPGGIPFILTGEFFQOSORPAAFIAGTVNWL 485  
 QY 461 SNFAVGLLFPFIQKSLDTCFLVFAFATCITGAILYFLVLPETKNTYAEISQAFSKENKA 520  
 DB 486 SNFAVGLLFPFIQKSLDTCFLVFAFATCITGAILYFLVLPETKNTYAEISQAFSKENKA 545  
 QY 521 YPPEEKIDSAVTD 533  
 DB 546 YPPEEKIDSAVTD 558  
 ID AAM78438 standard; protein; 408 AA.  
 XX AAM78438;  
 XX 06-NOV-2001 (first entry)  
 DE Human protein SEQ ID NO 1100.  
 XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 KW nervous system disorder; arthritis; inflammation.  
 XX Homo sapiens.  
 XX WO200157190-A2.  
 PD 09-AUG-2001.  
 XX 05-FEB-2001; 2001WO-US004098.  
 XX 03-FEB-2000; 2000US-00496914.  
 PR 27-APR-2000; 2000US-00560875.  
 PR 20-JUN-2000; 2000US-00598075.  
 PR 19-JUL-2000; 2000US-00820325.  
 PR 01-SEP-2000; 2000US-00654936.  
 PR 15-SEP-2000; 2000US-00663561.  
 PR 20-OCT-2000; 2000US-00693325.  
 PR 30-NOV-2000; 2000US-00728422.  
 XX (HYSE-) HYSEQ INC.  
 XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
 PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R;  
 XX WPI: 2001-476283/51.  
 DR N-PSDB; AAK51571.  
 XX Nucleic acids encoding polypeptides with cytokine-like activities, useful  
 PT in diagnosis and gene therapy.  
 XX Claim 20; Page 3326-3327; 622pp; English.  
 XX The invention relates to polynucleotides (AAK51456-AAK53435) and the  
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity relating to  
 CC cytokine, cell proliferation or cell differentiation or which may induce  
 CC production of other cytokines in other cell populations. The  
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
 CC peptide therapy. The polypeptides have various cytokine-like activities,  
 CC e.g. stem cell growth factor activity, haematopoiesis regulating  
 CC activity, tissue growth factor activity, immunomodulatory activity and  
 CC activin/inhibin activity and may be useful in the diagnosis and/or  
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
 CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
 CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
 CC sequence listing were missing at the time of publication

121 GLVGTLLIVKMIKGLVKRKHITLLANGFAISALLMACSLQAGAFEMLIYGRFIMGIDGGV 180
121 GLVGTLLIVKMIKGLVKRKHITLLANGFAISALLMACSLQAGAFEMLIYGRFIMGIDGGV 180
181 ALSVLPMYLSSEISPKIRGSLGQVTAIFTCIGVFTGQLLGLPELLGKESTWYLFVIVV 240
181 ALSVLPMYLSSEISPKIRGSLGQVTAIFTCIGVFTGQLLGLPELLGKESTWYLFVIVV 240
241 PAVVQLLSLPFLPDSRYLLEKHNARAVKAFQTFELGKADYSQEVVEVLAEHSVORSIR 300
241 PAVVQLLSLPFLDSDRYLLEKHNARAVKAFQTFELGKADYSQEVVEVLAEHSVORSIR 300
301 LVSLELLRAPYVRWQVTVITMACYQICGLNAINFYNTSIFGKAGISPAKIPYVTLST 360
301 LVSLELLRAPYVRWQVTVITMACYQICGLNAINFYNTSIFGKAGISPAKIPYVTLST 360
361 GGIETLAAVFSGLVIEHLGRPELLIGGFLMGLFFGTLITLTLQDHAPWVPLSIVGIL 420
361 GGIETLAAVFS----- 371
421 AIIASFCSGGGIPFILLTGSEFFQOSORPAAFIAGTWNWLSNFAVGLLPFFIOKSLDTYC 480
372 -----GIPFILLTGSEFFQOSORPAAFIAGTWNWLSNFAVGLLPFFIOKSLDTYC 420
481 FLVFAITCITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEKIDSAVTD 533
421 FLVFAITCITGAIYLYFVLPETKNRTYAEISQAFSKRNKAYPPEEKIDSAVTD 473
RESULT 11
ADC26253
ID ADC26253 standard; protein; 565 AA.
XX AC ADC26253;
XX AC ADC26253;
DT 18-DEC-2003 (first entry)
XX DE Human NOV17b protein.
XX KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;
XX KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;
XX KW transgenic; human.
XX OS Homo sapiens.
XX PN WO2003004687-A2.
XX PD 16-JAN-2003.
XX PF 03-JUL-2002; 2002WO-US021361.
XX PR 05-JUL-2001; 2001US-0303046P.
XX PR 09-JUL-2001; 2001US-0303828P.
XX PR 09-JUL-2001; 2001US-0304016P.
XX PR 11-JUL-2001; 2001US-0304502P.
XX PR 13-JUL-2001; 2001US-0305262P.
XX PR 16-JUL-2001; 2001US-0305673P.
XX PR 17-JUL-2001; 2001US-0306085P.
XX PR 24-JUL-2001; 2001US-0307536P.
XX PR 27-JUL-2001; 2001US-0308228P.
XX PR 30-JUL-2001; 2001US-0308877P.
XX PR 01-AUG-2001; 2001US-0309255P.
XX PR 17-AUG-2001; 2001US-0313328P.
XX PR 12-SEP-2001; 2001US-0318711P.
XX PR 19-SEP-2001; 2001US-0323380P.
XX PR 21-SEP-2001; 2001US-0323969P.
XX PR 04-JAN-2002; 2002US-0345022P.
XX PR 04-JAN-2002; 2002US-0345038P.
XX PR 28-FEB-2002; 2002US-0361172P.
XX PR 01-MAR-2002; 2002US-0360814P.
XX PR 01-MAR-2002; 2002US-0360830P.
XX PR 01-MAR-2002; 2002US-0361133P.
XX PR 01-MAR-2002; 2002US-0361147P.

05-JUL-2001; 2001US-0303046P.
09-JUL-2001; 2001US-0303828P.
09-JUL-2001; 2001US-0304016P.
11-JUL-2001; 2001US-0304502P.
13-JUL-2001; 2001US-0305262P.
16-JUL-2001; 2001US-0305673P.
17-JUL-2001; 2001US-0306085P.
24-JUL-2001; 2001US-0307536P.
27-JUL-2001; 2001US-0308228P.
30-JUL-2001; 2001US-0308877P.
01-AUG-2001; 2001US-0309255P.
17-AUG-2001; 2001US-0313328P.
12-SEP-2001; 2001US-0318711P.
19-SEP-2001; 2001US-0323380P.
21-SEP-2001; 2001US-0323969P.
04-JAN-2002; 2002US-0345022P.
04-JAN-2002; 2002US-0345038P.
28-FEB-2002; 2002US-0361172P.
01-MAR-2002; 2002US-0360814P.
01-MAR-2002; 2002US-0360830P.
01-MAR-2002; 2002US-0361133P.
01-MAR-2002; 2002US-0361147P.
05-MAR-2002; 2002US-0361677P.
02-APR-2002; 2002US-0363637P.
12-APR-2002; 2002US-0372326P.
16-APR-2002; 2002US-0372990P.
19-APR-2002; 2002US-0373881P.
19-APR-2002; 2002US-0373921P.
02-JUL-2002; 2002US-00188186.
(CURA-) CURAGEN CORP.
Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ,
Catterson DJ, Edinger S, Eissen AJ, Ellerman K, Gerlach V, Gorman L,
Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE,
Padigaru M, Patturajan M, Pena CE, Rastelli L, Shenoy S,
Shimkete RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ,
Vernet CAM, Voss EZ, Zhong M;
WPI; 2003-221607/21.
N-PSDB; ADC26254.
New isolated NOVX polypeptide, useful for determining the presence of, or
predisposition to a disease associated with altered levels of expression
of the polypeptide, and for treating or preventing cancer.
Claim 2; SEQ ID NO 80; 478pp; English.
The invention relates to a novel isolated NOV polypeptide. The
polypeptide of the invention demonstrates cytostatic activity and may be
used for determining the presence of, or predisposition to a disease
associated with altered levels of expression of the polypeptide,
including metabolic disorders, immune disorders, neurodegenerative
disorders, circulatory diseases, haemopoietic disorders, wasting diseases
and cancer. The polypeptide may also be utilised during gene therapy
procedures, vaccine development and transgenic animal production. The
current sequence is that of the human NOV protein of the invention.
Sequence 480 AA;
Query Match 80.6%; Score 2341; DB 7; Length 480;
Best Local Similarity 87.6%; Pred. No. 1e-229; 6; Indels 60; Gaps 1;
Matches 467; Conservative 0; Mismatches 6;
1 MARKQNRNSKELGLVPLTDTSHAGPPGGRALLECDHLRSGVPGRRKDWSCSLVAS 60
1 MARKQNRNSKELGLVPLTDTSHAGPPGGRALLECDHLRSGVPGRRKDWSCSLVAS 60
61 LAGAFGSSFLYGNLSVNNAPTYIKAFYNESWERRHGRPIDPDLTLTWSVTVSIFAIG 120
61 LAGAFGSPFLYGNLSVNNAPTYIKAFYNESWERRHGRPIDPDLTLTWSVTVSIFAIG 120





DR N-PSDB; AA160472.

XX Novel nucleic acids and polypeptides, useful for treating disorders such

PT as central nervous system injuries.

XX

PS Example 2; SEQ ID NO 6247; 10078pp; English.

XX

XX The invention relates to human nucleic acids (AA157798-AA161369) and the

CC encoded polypeptides (AA157798-AA161369) with neurotrophic,

CC immunosuppressive and cytostatic activity. The polynucleotides are useful

CC in gene therapy. A composition containing a polypeptide or polynucleotide

CC of the invention may be used to treat diseases of the peripheral nervous

CC system, such as peripheral nervous injuries, peripheral neuropathy and

CC localised neuropathies and central nervous system diseases, such as

CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic

CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the

CC utilisation of the activities such as: Immune system suppression,

CC Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic

CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,

CC assays for receptor activity, arthritis and inflammation, leukaemias and

CC C.N.S disorders. Note: The sequence data for this patent did not form

CC part of the printed specification

XX

XX Sequence 558 AA;

Qy Query Match 94.0%; Score 2730; DB 4; Length 558;

Db Best Local Similarity 99.6%; Pred. No. 2.4e-269;

Qy Matches 53; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 MAKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSYGPGRRKDKWCSLLVAS 60

19 MAKQNRNSKELGLVPLTDDTSHAGPPGGRALLECDHLRSYGPGRRKDKWCSLLVAS 78

61 LAGNFGSSFLYGNLSVNNAPTPVIKAFYNESERRHGRPIDDTTLTLLMSVTSIFAIG 120

79 LAGAFSSFLYGNLSVNNAPTPVIKAFYNESERRHGRPIDDTTLTLLMSVTSIFAIG 138

121 GLVGTLLVMKGVLRKHTLLANNFPAISAAALMACSLOAGAFEMLIIVGRFIMGIDGV 180

139 GLVGTLLVMKGVLRKHTLLANNFPAISAAALMACSLOAGAFEMLIIVGRFIMGIDGV 198

181 ALSVLPNYSISPEKIRSGLSQVTAIFICIGVFTGQLLGPBELLGHESTWPLFGVIV 240

199 ALSVLPNYSISPEKIRSGLSQVTAIFICIGVFTGQLLGPBELLGHESTWPLFGVIV 258

241 PAVVQLLSLPDPSRYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEASHVQSR 300

259 PAVVQLLSLPDPSRYLLEKNEARAVKAFOTFLGKADVSQVEEVLAEASHVQSR 318

301 LVSVEILLRABYVRQVTVVTVWACVQLCGLNALWFTNSIFGKAGIPPAKIPYVTLST 360

319 LVSVEILLRABYVRQVTVVTVWACVQLCGLNALWFTNSIFGKAGIPPAKIPYVTLST 378

361 GGIEFLAAVFGSLVIEHLGRPLLIIGFGLMGLFFGTITLTITLQDHAPWPVYLSIVGIL 420

379 GGIEFLAAVFGSLVIEHLGRPLLIIGFGLMGLFFGTITLTITLQDHAPWPVYLSIVGIL 438

421 AIIASFCSGPGGIPFILTGEFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 480

439 AIIASFCSGPGGIPFILTGEFFQSQRPAAFIAGTVNWLNSNFAVGLLFPFIQKSLDTC 498

481 FLVFATICTGAIVLYFVLPETKRTYAEISQAFSKRNKAYPPEKIDSATVD 533

499 FLVFATICTGAIVLYFVLPETKRTYAEISQAFSKRNKAYPPEKIDSATVD 551

RESULT 7

AAE16788

ID AAE16788 standard; protein; 537 AA.

XX

XX AAE16788;

AC

XX 09-APR-2002 (first entry)

DT

XX Human transporter and ion channel-25 (TRICH-25) protein.

XX

XX Human; transporter and ion channel-25; TRICH-25; neuroprotective; asthma;

XX neurotropic; cytostatic; cardiovascular; immunosuppressive; cardiomyopathy;

KW antiinflammatory; protein therapy; akinesia; cystic fibrosis; leukaemia;

KW Bell's palsy; amyotrophic lateral sclerosis; Alzheimer's disease; cancer;

KW amnesia; dementia; myocarditis; Duchenne's muscular dystrophy; AIDS;

KW Acquired Immune Deficiency Syndrome; Addison's disease; allergy; angina;

KW cell proliferative disorder; psoriasis; cardiac disease; hypertension;

KW bradyarrhythmia; gene expression; drug screening.

XX

OS Homo sapiens.

XX

XX Location/Qualifiers

Key 1..66

Peptide /label= Signal\_peptide

Domain 59..514

FT /note= "Sugar transporter domain"

Protein 67..537

FT /label= Mature\_TRICH\_25\_protein

Domain 112..128

FT /label= Transmembrane\_domain

Domain 173..198

FT /note= "Sugar transporter 2 motif"

Domain 371..386

FT /note= "Sugar transporter 1 motif"

Domain 385..404

FT /label= Transmembrane\_domain

Domain 414..436

FT /label= Transmembrane\_domain

Domain 479..497

FT /label= Transmembrane\_domain

XX

WO200192304-A2.

XX

06-DEC-2001.

XX

25-MAY-2001; 2001WO-US017065.

XX

26-MAY-2000; 2000US-0208424P.

XX

01-JUN-2000; 2000US-0209001P.

PR

08-JUN-2000; 2000US-0210588P.

PR

16-JUN-2000; 2000US-0212335P.

PR

22-JUN-2000; 2000US-0213747P.

PR

29-JUN-2000; 2000US-0215391P.

XX

(INCY-) INCYTE GENOMICS INC.

XX

Thornton M, Wallia NK, Yue H, Nguyen DB, Lai P, Gandhi AR;

PI Tribouley CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang Y;

PI Azimzai Y, Bruns CM, Griffin JA, Yang J, Sanjanwala MS, Raumann BE;

PI Lee BA, Hafalia A, Baughn MR, Green BD, Khan FA, Kearney L;

PI Elliot VS, Seilhamer JJ, Policky JL, Borowsky ML, Burford N, Ding L;

PI Lu DAM, Hillman JL;

XX

WPI; 2002-122055/16.

DR

N-PSDB; AAD27278.

XX

New human transporters and ion channels (TRICH) polypeptides useful for

PT diagnosing, treating or preventing disorders associated with aberrant

PT expression of TRICH.

XX

Claim 1; Page 182-184; 210pp; English.

PS

XX The invention relates to human transporters and ion channels (TRICH)

CC polypeptides and their cDNA molecules. The nucleic acid and polypeptide

CC sequences are useful in the diagnosis, treatment, and prevention of

CC disorders associated with transport (akinesia, cystic fibrosis, Bell's

CC palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,

CC amnesia, dementia); muscle (cardiomyopathy, myocarditis, Duchenne's

CC muscular dystrophy); immunological (AIDS, Addison's disease, allergies;

CC asthma); cell proliferative disorders (cancers, leukaemia, psoriasis);



Db 79 LAGAGSSFLYGYNLVSVNAPTPIYKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFATG 138  
QY 121 GLVGTLLIVKMGKVLGRKHTLLANNNGFAJSAALLMACSLQAGAFEMLIYGRFTMGIDGV 180  
Db 139 GLVGTLLIVKMGKVLGRKHTLLANNNGFAJSAALLMACSLQAGAFEMLIYGRFTMGIDGV 198  
QY 181 ALSVLPWYLSEISPKIRSGSLGQVTAIFICIGVFTQQLLGLPELLGKSTWPLYFVIVV 240  
Db 199 ALSVLPWYLSEISPKIRSGSLGQVTAIFICIGVFTQQLLGLPELLGKSTWPLYFVIVV 258  
QY 241 PAVVQLLSLPPFLDPSRYLLLEKHNARAVKAPQTFGLKADVSQEVVEVLAEHVORSIR 300  
Db 259 PAVVQLLSLPPFLDPSRYLLLEKHNARAVKAPQTFGLKADVSQEVVEVLAEHVORSIR 318  
QY 301 LVSVELELLRAPYRWQVTVITMACYQVLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 360  
Db 319 LVSVELELLRAPYRWQVTVITMACYQVLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 378  
QY 361 GGITETLAASFVGLVIEHLGRPLLIIGFGLMGLFFGTTLITTLQDHAPWVPLSIVGIL 420  
Db 379 GGITETLAASFVGLVIEHLGRPLLIIGFGLMGLFFGTTLITTLQDHAPWVPLSIVGIL 438  
QY 421 AIIASFCSGPGGIPFILTGEFFQSQSPRAAFIAGTVNMLSNFAVGLLPFFIQKSLDTYC 480  
Db 439 AIIASFCSGPGGIPFILTGEFFQSQSPRAAFIAGTVNMLSNFAVGLLPFFIQKSLDTYC 498  
QY 481 FLVPATCITGAILYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 533  
Db 499 FLVPATCITGAILYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTD 551  
RESULT 6  
AAM41316  
ID AAM41316 standard; protein; 558 AA.  
XX AC AAM41316;  
XX DT 22-OCT-2001 (first entry)  
XX DE Human polypeptide SEQ ID NO 6247.  
XX KW Human; nontropic; immunosuppressant; cytostatic; gene therapy; cancer;  
KW peripheral nervous system; neuropathy; central nervous system; CNS;  
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
KW leukaemia.  
XX OS Homo sapiens.  
XX PN WO200153312-A1.  
XX PD 26-JUL-2001.  
XX PF 26-DEC-2000; 2000WO-US034263.  
XX PR 23-DEC-1999; 99US-00471275.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 20-JUN-2000; 2000US-00598042.  
XX PR 19-JUL-2000; 2000US-00620312.  
XX PR 03-AUG-2000; 2000US-00653450.  
XX PR 14-SEP-2000; 2000US-00662191.  
XX PR 19-OCT-2000; 2000US-00693036.  
XX PR 29-NOV-2000; 2000US-00727344.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;  
PI Zhou P, Goodrich R, Drmanac RT;  
XX WPI; 2001-442253/47.  
DR

RESULT 5  
AAM79422  
ID AAM79422 standard; protein; 558 AA.  
XX AC AAM79422;  
XX DT 06-NOV-2001 (first entry)  
XX DE Human protein SEQ ID NO 3068.  
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;  
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
KW nervous system disorder; arthritis; inflammation.  
XX OS Homo sapiens.  
XX PN WO200157190-A2.  
XX PD 09-AUG-2001.  
XX PF 05-FEB-2001; 2001WO-US004098.  
XX PR 03-FEB-2000; 2000US-00496914.  
XX PR 27-APR-2000; 2000US-00560875.  
XX PR 20-JUN-2000; 2000US-00598075.  
XX PR 19-JUL-2000; 2000US-00620325.  
XX PR 01-SEP-2000; 2000US-00654936.  
XX PR 15-SEP-2000; 2000US-00663561.  
XX PR 20-OCT-2000; 2000US-00693325.  
XX PR 30-NOV-2000; 2000US-00728422.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;  
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;  
PI Xue AJ, Yang Y, Wehrman T, Goodrich R;  
XX WPI; 2001-476283/51.  
XX DR N-PSDB; AAK52555.  
XX KW Nucleic acids encoding polypeptides with cytokine-like activities, useful  
PT in diagnosis and gene therapy.  
XX PS Claim 20; Page 242; 6221pp; English.  
XX CC The invention relates to polynucleotides (AAK51456-AAK53435) and the  
CC encoded polypeptides (AAM78323-AAW80302) that exhibit activity elating to  
CC cytokine, cell proliferation or cell differentiation or which may induce  
CC production of other cytokines in other cell populations. The  
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or  
CC peptide therapy. The polypeptides have various cytokine-like activities,  
CC e.g. stem cell growth factor activity, haematopoiesis regulating  
CC activity, tissue growth factor activity, immunomodulatory activity and  
CC activin/inhibin activity and may be useful in the diagnosis and/or  
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and  
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111  
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the  
CC sequence listing were missing at the time of publication  
XX SQ Sequence 558 AA;  
Query Match 94.3%; Score 2737; DB 4; Length 558;  
Best Local Similarity 99.8%; Pred. No. 4.7e-270;  
Matches 532; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRALLECDHLRSVPGRRKDKWSCSILLVAS 60  
Db 19 MARKQNRNSKELGLVPLTDDTSHAGPPGPGRALLECDHLRSVPGRRKDKWSCSILLVAS 78  
QY 61 LAGAGSSFLYGYNLVSVNAPTPIYKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFATG 120



[illegible]

XX (MILL-) MILLENNIUM PHARM INC.  
 XX Tartaglia LA, Weng X;  
 PI WPI; 1999-526192/44.  
 XX N-PSDB; AAZ11502.  
 DR A novel mammalian transporter protein and the gene encoding it, are  
 XX useful in the diagnosis and treatment of disorders associated with  
 PT aberrant sugar transport.  
 PT Claim 1; Fig 2A-D; 48pp; English.  
 XX This represents a human glucose transporter (GLUTX) protein. The protein  
 XX can be expressed by standard recombinant methodology. GLUTX nucleic acids  
 CC are useful as hybridization probes for detecting the presence of GLUTX  
 CC DNA in a sample, useful for diagnosing conditions associated with  
 CC aberrant expression levels of GLUTX. The GLUTX gene is also useful as a  
 CC therapeutic agent for regulating translation of GLUTX mRNA, and for  
 CC treatment of disorders associated with aberrant expression of GLUTX and  
 CC aberrant hexose transport. It is useful for generating GLUTX specific  
 CC antibodies, identifying agonists and antagonists of GLUTX, and  
 CC identifying nucleic acids in other species encoding nucleic acids  
 CC homologous to GLUTX. GLUTX nucleic acids are also useful for identifying  
 CC the chromosomal location of GLUTX, and as tissue specific markers  
 XX Sequence 563 AA;  
 SQ Query Match 100.0%; Score 2903; DB 2; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 5.3e-287;  
 Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MARKQNSKELGLVPLTDDTSHAGPPGPRALLECDHLRSGVPGRRKDWSCSLIVAS 60  
 DB 1 MARKQNSKELGLVPLTDDTSHAGPPGPRALLECDHLRSGVPGRRKDWSCSLIVAS 60  
 QY 61 LAGAFGSLGYNLVSVNAPTPYKAFYNESWERRGRPIDPTLLMSVTVSIFAIG 120  
 DB 61 LAGAFGSLGYNLVSVNAPTPYKAFYNESWERRGRPIDPTLLMSVTVSIFAIG 120  
 QY 121 GLVGTLLVMIGKVLGRKHTLLANGFAISAAALMACSLQAGAFEMLIIVGRFIMGIDGV 180  
 DB 121 GLVGTLLVMIGKVLGRKHTLLANGFAISAAALMACSLQAGAFEMLIIVGRFIMGIDGV 180  
 QY 181 ALSVLPWYLSISPKRIGSLGQVTAIFCIGVFTGQLLGPPELLLGKSTWPLYFGVIV 240  
 DB 181 ALSVLPWYLSISPKRIGSLGQVTAIFCIGVFTGQLLGPPELLLGKSTWPLYFGVIV 240  
 QY 241 PAVVQLLSLPDPSRYLLEKNEARAVKAFOTFLGKADVSOEVEVLAEASHVQSR 300  
 DB 241 PAVVQLLSLPDPSRYLLEKNEARAVKAFOTFLGKADVSOEVEVLAEASHVQSR 300  
 QY 301 LVSVELELRAPYVRWQVTVITWACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 360  
 DB 301 LVSVELELRAPYVRWQVTVITWACYQLCGLNAINFYTNISIFGKAGIPPAKIPYVTLST 360  
 QY 361 GGIETLAAVFSGLVIEHGRPLLGFGMLGFFGLTITLTLQDHAPWVPLYSIVGIL 420  
 DB 361 GGIETLAAVFSGLVIEHGRPLLGFGMLGFFGLTITLTLQDHAPWVPLYSIVGIL 420  
 QY 421 AIIASFCSGPGIPILTGEPFQSQRPDAPIIAGTVNLSNPAVGLLFPPIQKSLTYC 480  
 DB 421 AIIASFCSGPGIPILTGEPFQSQRPDAPIIAGTVNLSNPAVGLLFPPIQKSLTYC 480  
 QY 481 FLVPTATICTGAIVLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTDAPASPP 540  
 DB 481 FLVPTATICTGAIVLYFVLPETKNTYAEISQAFSKRNKAYPPEEKIDSAVTDAPASPP 540  
 QY 541 TTPNTANIQAAATTTATKKEHPL 563  
 DB 541 TTPNTANIQAAATTTATKKEHPL 563

RESULT 2  
 AAB30521  
 ID AAB30521 standard; protein; 563 AA.  
 XX  
 AC AAB30521;  
 XX  
 DT 06-MAR-2001 (first entry)  
 XX  
 DE Amino acid sequence of a human GLUTX polypeptide.  
 XX  
 KW Human; GLUTX; glucose transporter; hexose transport;  
 KW tissue-specific marker.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 52..71  
 FT Domain /note= "transmembrane domain"  
 FT Domain 108..128  
 FT Domain /note= "transmembrane domain"  
 FT Domain 141..159  
 FT Domain /note= "transmembrane domain"  
 FT Domain 166..189  
 FT Domain /note= "transmembrane domain"  
 FT Domain 204..221  
 FT Domain /note= "transmembrane domain"  
 FT Domain 233..252  
 FT Domain /note= "transmembrane domain"  
 FT Domain 317..333  
 FT Domain /note= "transmembrane domain"  
 FT Domain 355..375  
 FT Domain /note= "transmembrane domain"  
 FT Domain 383..404  
 FT Domain /note= "transmembrane domain"  
 FT Domain 413..437  
 FT Domain /note= "transmembrane domain"  
 FT Domain 449..472  
 FT Domain /note= "transmembrane domain"  
 FT Domain 481..499  
 FT Domain /note= "transmembrane domain"  
 XX  
 FN US6136547-A.  
 XX  
 PD 24-OCT-2000.  
 XX  
 PF 26-APR-1999; 99US-00299549.  
 XX  
 PR 26-FEB-1998; 98US-00031392.  
 XX  
 PA (MILL-) MILLENNIUM PHARM INC.  
 XX  
 PI Weng X, Tartaglia LA;  
 XX  
 DR WPI; 2001-040236/05.  
 DR N-PSDB; AAC62143.  
 XX  
 PT New human glucose transporter protein and nucleic acid encoding the  
 PT protein, useful in the diagnosis and treatment of disorders associated  
 PT with aberrant hexose transport.  
 XX  
 PS Claim 1; Fig 2A-D; 48pp; English.  
 XX  
 CC The present sequence represents a human GLUTX polypeptide. GLUTX is a  
 CC glucose transporter protein. The GLUTX polypeptides and polynucleotides  
 CC are useful in the diagnosis and treatment of disorders associated with  
 CC aberrant hexose transport. The GLUTX polypeptide is useful for  
 CC determining whether a given disorder is associated with aberrant  
 CC expression of GLUTX or activity of GLUTX. The nucleic acid molecule and  
 CC the GLUTX polypeptide are useful as diagnostic or therapeutic agents, or  
 CC they can be used to generate antibodies or identify small molecules that,  
 CC in turn are clinically useful. The GLUTX nucleic acid molecules are  
 CC useful for identifying the chromosomal location of GLUTX and as tissue-

OM protein - protein search, using sw model

Run on: April 6, 2004, 09:59:52 ; Search time 59 Seconds  
(without alignments)  
2696.174 Million cell updates/sec

Title: US-09-981-947B-2  
Perfect score: 2903  
Sequence: 1 MARKQNRNSKELGLVPLTDD.....NTAWIQAAATTTATKEHPL 563

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq 25Jan04: \*  
1: geneseqp1980s: \*  
2: geneseqp1990s: \*  
3: geneseqp2000s: \*  
4: geneseqp2001s: \*  
5: geneseqp2002s: \*  
6: geneseqp2003as: \*  
7: geneseqp2003bs: \*  
8: geneseqp2004s: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2903	100.0	563	2	AAV27286 Human glu
2	2903	100.0	563	4	AB30521 Amino aci
3	2903	100.0	563	5	AAE14611 Human glu
4	2737	94.3	558	4	ABB11910 Human glu
5	2737	94.3	558	4	AAW79422 Human pro
6	2720	94.0	558	4	AAW41116 Human pol
7	2723	93.8	537	5	AAE16788 Human tra
8	2527.5	87.1	507	7	ADC26251 Human NOV
9	2396	82.5	472	5	ABP52206 Human fac
10	2341	80.6	480	7	ADC26255 Human NOV
11	2337.5	80.5	565	7	ADC26253 Human NOV
12	2018	69.5	408	4	AAW78438 Human pro
13	2017	69.5	422	7	ADC26285 Human NOV
14	2012	69.3	408	4	AAW39530 Human pol
15	1939	66.8	416	4	AAE04893 Human tra
16	1929	66.4	396	7	ADC26261 Human NOV
17	1920	66.1	396	7	ADC26259 Human NOV
18	1911	65.8	396	7	ADC26283 Human NOV
19	1721	59.3	362	7	ADC26257 Human NOV
20	1122.5	38.7	501	5	ABP65173 Hypoxia-r
21	1122.5	38.7	501	6	ABP65173 Hypoxia-r
22	1122.5	38.7	501	7	ABP58597 Human can
23	1122.5	38.7	501	7	ABP56199 Human glu
24	1119.5	38.6	512	5	AAO14130 Human tra
25	1119.5	38.6	512	7	ABP56198 Human glu

26	1119.5	38.6	512	7	ADD01365	Add01365 Human glu
27	1117	38.5	521	7	ADD01387	Add01387 Human var
28	1108	38.2	500	2	AAV27291	AAV27291 Glucose t
29	1081	37.2	493	4	AAE30522	AAE30522 Glucose t
30	1081	37.2	493	4	AAE30522	AAE30522 Amino aci
31	1073	37.0	565	6	AAE35305	AAE35305 Human sbg
32	1001.5	34.5	500	5	ABP80602	ABP80602 Human sbg
33	1000.5	34.5	499	5	AAO14198	AAO14198 Human tra
34	999	34.4	496	5	AAE88405	AAE88405 Human mem
35	999	34.4	496	5	ABP80601	ABP80601 Human sbg
36	997.5	34.4	483	5	ABG70393	ABG70393 Human glu
37	979	33.7	524	6	ABG70513	ABG70513 Human tra
38	952.5	32.8	486	6	ABO00640	ABO00640 Novel hum
39	936.5	32.3	492	2	AAW17835	AAW17835 Human glu
40	936.5	32.3	492	5	ABP65129	ABP65129 Hypoxia-r
41	936.5	32.3	492	6	ABU56445	ABU56445 Lung canc
42	936	32.2	471	5	ABG70347	ABG70347 Novel hum
43	935	32.2	501	5	ABG70348	ABG70348 Novel hum
44	934.5	32.2	480	6	ABJ37062	ABJ37062 Human bre
45	934.5	32.2	492	2	AAE11360	AAE11360 Glucose T

ALIGNMENTS

RESULT 1  
AAV27286  
ID AAV27286 standard; protein; 563 AA.  
XX  
AC AAV27286;  
XX  
DT 05-NOV-1999 (first entry)  
XX  
DE Human glucose transporter protein GLUTX.  
XX  
KW Human; glucose transporter; GLUTX; hexose transport; gene therapy;  
KW chromosome mapping.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Domain 52..71  
FT /note= "transmembrane domain 1"  
FT Domain 108..128  
FT /note= "transmembrane domain 2"  
FT Domain 141..159  
FT /note= "transmembrane domain 3"  
FT Domain 166..189  
FT /note= "transmembrane domain 4"  
FT Domain 204..221  
FT /note= "transmembrane domain 5"  
FT Domain 233..252  
FT /note= "transmembrane domain 6"  
FT Domain 317..333  
FT /note= "transmembrane domain 7"  
FT Domain 355..375  
FT /note= "transmembrane domain 8"  
FT Domain 383..404  
FT /note= "transmembrane domain 9"  
FT Domain 413..437  
FT /note= "transmembrane domain 10"  
FT Domain 449..472  
FT /note= "transmembrane domain 11"  
FT Domain 481..499  
FT /note= "transmembrane domain 12"  
XX  
US5942398-A.  
XX  
PD 24-AUG-1999.  
XX  
PF 26-FEB-1998;  
XX  
PR 26-FEB-1998;  
98US-00031392.  
98US-00031392.

XX PA (MILL-) MILLENNIUM PHARM INC.  
XX PI Tartaglia LA, Weng X;  
XX XX WPI; 1999-526192/44.  
DR N-PSDB; AAC211502.  
XX XX A novel mammalian transporter protein and the gene encoding it, are  
PT useful in the diagnosis and treatment of disorders associated with  
PT aberrant sugar transport.  
XX PS Claim 1; Fig 2A-D; 48pp; English.  
XX CC This represents a human glucose transporter (GLUTX) protein. The protein  
CC can be expressed by standard recombinant methodology. GLUTX nucleic acids  
CC are useful as hybridization probes for detecting the presence of GLUTX  
CC DNA in a sample, useful for diagnosing conditions associated with  
CC aberrant expression levels of GLUTX. The GLUTX gene is also useful as a  
CC therapeutic agent for regulating translation of GLUTX mRNA, and for  
CC treatment of disorders associated with aberrant expression of GLUTX and  
CC aberrant hexose transport. It is useful for generating GLUTX specific  
CC antibodies, identifying agonists and antagonists of GLUTX, and  
CC identifying nucleic acids in other species encoding nucleic acids  
CC homologous to GLUTX. GLUTX nucleic acids are also useful for identifying  
CC the chromosomal location of GLUTX, and as tissue specific markers  
XX SQ Sequence 563 AA;

Query Match 100.0%; Score 2903; DB 2; Length 563;  
Best Local Similarity 100.0%; Pred. No. 5.3e-287;  
Matches 563; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPRALLSCDHLRSVPGRRRKQWSCSLLVAS 60  
Db 1 MARKQNRNSKELGLVPLTDDTSHAGPPGPRALLSCDHLRSVPGRRRKQWSCSLLVAS 60  
QY 61 LAGAFSSFLYGNLSVNVNAPYIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 120  
Db 61 LAGAFSSFLYGNLSVNVNAPYIKAFYNESWERRHGRPIDPDTLLTLLWSVTVSIFAIG 120  
QY 121 GLVGTLLVIMGKVLGRKHTLLANNGFAISALLMACSLQAGAFENLVGRFIMGDGV 180  
Db 121 GLVGTLLVIMGKVLGRKHTLLANNGFAISALLMACSLQAGAFENLVGRFIMGDGV 180  
QY 181 ALSVLPWYLSEISPKIRSGLOVTAIFICIGVFTQGLLGLPGLLKGSTWYLFQVIV 240  
Db 181 ALSVLPWYLSEISPKIRSGLOVTAIFICIGVFTQGLLGLPGLLKGSTWYLFQVIV 240  
QY 241 PAVVQLLSLPLDPSRYYLLEKHNARAVKAFOTELGKADYSQVEEVLAEHVQSRIR 300  
Db 241 PAVVQLLSLPLDPSRYYLLEKHNARAVKAFOTELGKADYSQVEEVLAEHVQSRIR 300  
QY 301 LVSLELLRAPYVRWQVTVITMACYQCLGNALWFTYNSIFGKAGIPPAKIPYVTLST 360  
Db 301 LVSLELLRAPYVRWQVTVITMACYQCLGNALWFTYNSIFGKAGIPPAKIPYVTLST 360  
QY 361 GGIETLAAPVSGLVIEHLGRRLPILGGFGLMGLFFGTLLITLTDQHPAPWYLSIVGIL 420  
Db 361 GGIETLAAPVSGLVIEHLGRRLPILGGFGLMGLFFGTLLITLTDQHPAPWYLSIVGIL 420  
QY 421 AIIASPCSGPGIPFILTGEFFQOSQORPAFIAGTNWLSNPAVGLLFPFQKSLDTC 480  
Db 421 AIIASPCSGPGIPFILTGEFFQOSQORPAFIAGTNWLSNPAVGLLFPFQKSLDTC 480  
QY 481 FLVPATICIGALYLVFVLPETKNTVAETISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540  
Db 481 FLVPATICIGALYLVFVLPETKNTVAETISQAFSKRNKAYPPEEKIDSAVTDAPASSPF 540  
QY 541 TTPNTAWIQAAATTTATKKEHPL 563  
Db 541 TTPNTAWIQAAATTTATKKEHPL 563

RESULT 2  
AAB30521  
ID AAB30521 standard; protein; 563 AA.  
XX AC AAB30521;  
XX XX 06-MAR-2001 (first entry)  
XX XX Amino acid sequence of a human GLUTX polypeptide.  
XX XX Human; GLUTX; glucose transporter; hexose transport;  
XX XX tissue-specific marker.  
XX XX Homo sapiens.  
XX PH Key Location/Qualifiers  
XX FT Domain 52..71 /note= "transmembrane domain"  
XX FT Domain 108..128 /note= "transmembrane domain"  
XX FT Domain 141..159 /note= "transmembrane domain"  
XX FT Domain 166..189 /note= "transmembrane domain"  
XX FT Domain 204..221 /note= "transmembrane domain"  
XX FT Domain 233..252 /note= "transmembrane domain"  
XX FT Domain 317..333 /note= "transmembrane domain"  
XX FT Domain 355..375 /note= "transmembrane domain"  
XX FT Domain 383..404 /note= "transmembrane domain"  
XX FT Domain 413..437 /note= "transmembrane domain"  
XX FT Domain 449..472 /note= "transmembrane domain"  
XX FT Domain 481..499 /note= "transmembrane domain"  
XX FT US6136547-A.  
XX PD 24-OCT-2000.  
XX PF 26-APR-1999; 99US-00299549.  
XX PR 26-FEB-1998; 98US-00031392.  
XX PA (MILL-) MILLENNIUM PHARM INC.  
XX XX Weng X, Tartaglia LA;  
XX XX WPI; 2001-040236/05.  
XX DR N-PSDB; AAC62143.  
XX XX New human glucose transporter protein and nucleic acid encoding the  
PT protein, useful in the diagnosis and treatment of disorders associated  
PT with aberrant hexose transport.  
XX PS Claim 1; Fig 2A-D; 48pp; English.  
XX CC The present sequence represents a human GLUTX polypeptide. GLUTX is a  
CC glucose transporter protein. The GLUTX polypeptides and polynucleotides  
CC are useful in the diagnosis and treatment of disorders associated with  
CC aberrant hexose transport. The GLUTX polypeptide is useful for  
CC determining whether a given disorder is associated with aberrant  
CC expression of GLUTX or activity of GLUTX. The nucleic acid molecule and  
CC the GLUTX polypeptide are useful as diagnostic or therapeutic agents, or  
CC they can be used to generate antibodies or identify small molecules that  
CC in turn are clinically useful. The GLUTX nucleic acid molecules are  
CC useful for identifying the chromosomal location of GLUTX and as tissue-

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

QM protein - protein search, using sw model

Run on: April 6, 2004, 09:59:52 ; Search time 59 seconds

(without alignments)  
2696.174 Million cell updates/sec

Title: US-09-981-947B-2

Perfect score: 2903

Sequence: 1 MARKQNRNSKELGLVPLTDD.....NTAWIQARATTTATKKEHPL 563

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*

1: Geneseqp1980s.\*

2: Geneseqp1990s.\*

3: Geneseqp2000s.\*

4: Geneseqp2001s.\*

5: Geneseqp2002s.\*

6: Geneseqp2003as.\*

7: Geneseqp2003bs.\*

8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2903	100.0	563	2	AAY27286 Human glu
2	2903	100.0	563	4	AAB30521 Amino aci
3	2903	100.0	563	5	AAE14611 Human glu
4	2737	94.3	558	4	ABB11910 Human GLU
5	2737	94.3	558	4	AAM79422 Human pro
6	2730	94.0	558	4	AAM41316 Human pol
7	2723	93.8	537	5	AAE16788 Human tra
8	2597.5	87.1	507	7	ADC26251 Human NOV
9	2396	82.5	472	5	ABP52206 Human fac
10	2341	80.6	480	7	ADC26255 Human NOV
11	2337.5	80.5	565	7	ADC26253 Human NOV
12	2018	69.5	408	4	AAM78438 Human pro
13	2017	69.5	422	7	ADC26265 Human NOV
14	2012	69.3	408	4	AAM39530 Human tra
15	1939	66.8	416	4	AAE04893 Human pol
16	1929	65.4	396	7	ADC26261 Human NOV
17	1920	66.1	396	7	ADC26259 Human NOV
18	1911	65.8	396	7	ADC26263 Human NOV
19	1721	59.3	362	7	ADC26257 Human NOV
20	1122.5	38.7	501	6	ABP65173 Hypoxia-x
21	1122.5	38.7	501	6	ABR58597 Human can
22	1122.5	38.7	501	7	ABR56199 Human glu
23	1122.5	38.7	501	7	ADD01473 Human GLU
24	1119.5	38.6	512	5	AAO14190 Human tra
25	1119.5	38.6	512	7	ABR56198 Human glu

26	1119.5	38.6	512	7	ADD01365	Add01365 Human glu
27	1117	38.5	521	7	ADD01387	Add01387 Human var
28	1108	38.2	500	2	AAY27291	Aay27291 Glucose t
29	1081	37.2	493	2	AAY27292	Aay27292 Glucose t
30	1081	37.2	493	4	AAB30522	Aab30522 Amino aci
31	1073	37.0	565	6	AAE33305	Aae33305 Human 637
32	1001.5	34.5	500	5	ABB80602	Abb80602 Human sbg
33	1000.5	34.5	499	5	AAO14198	Aao14198 Human tra
34	999	34.4	496	4	AAE88405	Aae88405 Human mem
35	999	34.4	496	5	ABB80601	Abb80601 Human sbg
36	997.5	34.4	483	5	ABG70393	Abg70393 Human glu
37	979	33.7	524	6	ABG70513	Abg70513 Human 324
38	952.5	32.8	486	6	ABO00640	Aboc00640 Novel hum
39	936.5	32.3	492	2	AAW17835	Aaw17835 Human glu
40	936.5	32.3	492	5	ABP65129	Abp65129 Hypoxia-x
41	936.5	32.3	492	6	ABU56445	Abu56445 Lung canc
42	936	32.2	471	5	ABG70347	Abg70347 Novel hum
43	935	32.2	501	5	ABG70348	Abg70348 Novel hum
44	934.5	32.2	480	6	ABJ37062	Abj37062 Human bre
45	934.5	32.2	492	2	AAI11360	Aai11360 Glucose T

#### ALIGNMENTS

##### RESULT 1

AAAY27286

ID AAY27286 standard; protein; 563 AA.

XX

AC AAY27286;

XX

DT 05-NOV-1999 (first entry)

XX

DE Human glucose transporter protein GLUTX.

XX

KW Human; glucose transporter; GLUTX; hexose transport; gene therapy; chromosome mapping.

KW

XX Homo sapiens.

OS

Key Location/Qualifiers

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain

FT Domain



```

BM925071      1297 bp      mRNA      linear      EST 12-MAR-2002
AGENCOURT 6627390 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5762360
5', mRNA sequence.
BM925071
BM925071.1 GI:19375450
EST.
Homo sapiens (human)
Homo sapiens
Eumariyota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1297)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12812 row: h column: 09
High quality sequence start: 3
High quality sequence stop: 436.
Location/Qualifiers
1. 1297
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5762360"
/lab_host="DH10B"
/clone_lib="NIH_MGC_122"
/note="Organ: pooled lung and spleen; Vector: pCMV-SPORTS-6;
Site 1: NotI; Site 2: EcoRV (destroyed); RNA source
anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dt
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH MGC Library."
FEATURES
source

```

ORIGIN	Query Match	19.3%;	Score 452.8;	DB 12;	Length 1297;
	Best Local Similarity	95.9%;	Pred. No. 5.1e-53;		
	Matches 487;	Conservative	0;	Mismatches 17;	Indels 4; Gaps 2;
QY	17	GCCTTGGCAGAGCTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG	76		
Db	27	GCCTTGGCAGAGCTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG	86		
QY	77	CAAGGAACAANAATAGGAATTCGAAGAACTGGGCCTAGTTCCTCTCAGAGATGACACCA	136		
Db	87	CAAGGAACAANAATAGGAATTCGAAGAACTGGGCCTAGTTCCTCTCAGAGATGACACCA	146		
QY	137	GCCACGCCGGGCTCCAGGGCCAGGGGGCACTGCTGGAGTCTGACCACTTGAGGAGTG	196		
Db	147	GCCACGCCAGGCTCCAGGGCCAGGGGGCACTGCTGGAGTCTGACCACTTGAGGAGTG	206		
QY	197	GGGTGCCAGGTGGAAGGAGAAGAAGACTGGTCCCTCGCTCCCTCGTGGGCTCCCTCG	256		
Db	207	GGGTGCCAGGTGGAAGGAGAAGAAGACTGGTCCCTCGCTCCCTCGTGGGCTCCCTCG	266		
QY	257	CGSGCGCCCTTGGCTCCCTCCCTCTACGGCTACAACCTGTCTCGGTGGTGAATGCCCCCA	316		
Db	267	CGSGCGCCCTTGGCTCCCTCCCTCTACGGCTACAACCTGTCTCGGTGGTGAATGCCCCCA	326		
QY	317	CCCGGTACATCAAGGCGCTTTTACAATGAGTCAATGGGAAAGAAGGCATGGACGTCCAATAG	376		

D <sub>b</sub>	327	CCCCGTACATCAAGCGCTTTTACAATGAGTCATGGGAAGAAGGCATTGCACCTCCCAATAG	386
Q <sub>y</sub>	377	ACCAGACACTCTGACTCTGCTCTGCTGCTGCTCATTTGCCATTCGGTGGAC	436
D <sub>b</sub>	387	ACCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTCATAATTCGCCATTCGGTGGAC	446
Q <sub>y</sub>	437	TTCGGGGACGTTAA-TTGTAAGATGATGGAAGGTTCTTCGGGAGG--TAGCACACT	492
D <sub>b</sub>	447	TTCGGGGACATTAACTCTGAAGATGATTGGAAGGTTCTTCGGGAGGAAACACACCTT	506
Q <sub>y</sub>	493	TTCGTGGCCCAATAATGGGHTTGCATTT	520
D <sub>b</sub>	507	TGGTGGGCCCAATTAATGGGGTTTGCATTT	534

RESULT 15

BE745805	601590005F1 NTH_MGC_9 Homo sapiens cDNA clone IMAGE:3928801 5',	961 bp	mRNA	linear	EST 15-SEP-2000			
LOCUS	mRNA sequence.							
DEFINITION	mRNA sequence.							
ACCESSION	BE745805							
VERSION	BE745805.1	GI:10159797						
KEYWORDS	EST.							
SOURCE	Homo sapiens (human)							
ORGANISM	Homo sapiens							
	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;							
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
REFERENCE	1 (bases 1 to 961)							
AUTHORS	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .							
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)							
JOURNAL	Unpublished (1999)							
COMMENT	Contact: Robert Strausberg, Ph.D.							

Email: [cgapbs@mail.nih.gov](mailto:cgapbs@mail.nih.gov)  
Tissue Procurement: DCTD/DTP  
CDNA Library Preparation: Ling Hong/Rubin Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Incyte Genomics, Inc  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: [image.llnl.gov](http://image.llnl.gov)  
plate: L10CM760 row: j column: 02  
high quality sequence stop: 752.

[illegible]

```

1. .961
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3928801"
/tissue_type="adenocarcinoma cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 9"
/notes="Organ: ovary; Vector: pOTE7; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally
cloned into EcoRI/XhoI sites using the following 5'
adaptor: GGACACGAG(G). Size-selected >500bp for average
insert size 1.8kb. Library constructed by Ling Hong in
the laboratory of Gerald M. Rubin (University of
California, Berkeley) using ZAP-cDNA synthesis kit
(Stratagene, La Jolla, CA) and electroporated into DH10B
cells."

```

## ORIGIN

Query Match	18.8%;	Score 439.4;	DB 10;	Length 961;
Best Local Similarity	97.2%;	Pred. No. 46-51;		
Matches	479;	Conservative	0;	Mismatches 11;
				Indels 3;
				Gaps 3;

QY 217 AGAAGGAC TGGTCTGCTCGTCTCCTCGTGGCTCCCTCGCGGGCCCTTCGGCTCCTCC 276

D5	274	AAATTGGACATGCTCCAGCCTCGTTCCTCGTGCCCCTGCCTCGTGGGTGAATGCC	CCCCCACCCCTCATCATCAAGGCGTTT	336	
Ov	277	TTCCTCTA	GCGCTACAACCTGTCGGTGGTGAATGCC	CCCCCACCCCTCATCATCAAGGCGTTT	336

334 TTCTCTACGGGTACACCTGTCTGGTGTGAATGCCCCACCCCGTACATCAAGGCCTTT 393

337 TACAATCAGTCATGGGAAGAAGGCATGGACGTCCATAGACCCAGACACTCTGACTCTG 396

Matches 587; Conservative 0; Mismatches 112; Indels 1; Gaps 1;  
Qy 950 GCACGTGCGAGAGAGATCCGGCTGTGTCTGCTGTGAGCTGCTGAGAGTCCCTACG 1009  
Db 1 GCGCGGTGCAAGGAACCTCCGGCTGTGTCTGCTGTGAGCTGCTGAGGCAACCCCTTG 60  
Qy 1010 TCCGCTGGCAGGTGGTCAACGCTGATGCTCAACATGCGCTCTACCACTGCTGGCTCA 1069  
Db 61 TCCGCTGGCAGGTGATCACCCTGATATATACCACTGCGCAGCTACCACTATGAGGACTCA 120  
Qy 1070 ATGCAATTTGGTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCCGCAAGA 1129  
Db 121 ATGGATCTGGTCTACCAACAGCATCTTTGGAAAGCTGGGATCCCTCAGGACAAGA 180  
Qy 1130 TCCCATACGTCACTTCACTGAGTACAGGGGCACTGAGACTTTGGTGGCGTCTTCTGGTT 1189  
Db 181 TCCCATACATCACTTCACTGAGTACAGGGGCACTGAGACTTTGGTGGCGTCTTCTGGTT 240  
Qy 1190 TGGTCACTGAGCACTGGGAGGAGACCTCTCTCATCTGCTGCTTTGGCTCATGGCC 1249  
Db 241 TGGTCACTGAGCGTTAGGAGGAGACCTCTCTCATCTGCTGCTTTGGCTCATGGCC 300  
Qy 1250 TCTCTTTGGAGCCCTCACCATCAGCTGACCTGAGGACCAAGCCCTGGGCTCCCT 1309  
Db 301 TCTCTTTGGAGCCCTCACCAGCACCTGAGGACCAAGCTGAGGCTCCCTGGGCTCCCT 360  
Qy 1310 ACCTGAGTATGCTGGGCACTTGGCCATCATCGCTCTTCTGAGTGGGCAAGGTGGGA 1369  
Db 361 ACCTGAGTATGCTGGGCACTTGGCCATCATCGCTCTTCTGAGTGGGCAAGGTGGGA 420  
Qy 1370 TCCGTTTCACTTGAAGTGGTGGTTCCTCAGCAATCTCAGCGCGGCTGCTTCAATCA 1429  
Db 421 TCCGTTTCACTTGAAGTGGTGGTTCCTCAGCAATCTCAGCGCGGCTGCTTCAATCA 480  
Qy 1430 TTGAGGACCGTCAACTGCTGCTCAACTTGGCTGCTGCTGCTGCTTCCCTTCCATTC 1489  
Db 481 TCGCAGGACAGTCAACTGCTGCTCAACTTGGCTGCTGCTGCTTCCCTTCCATTC 540  
Qy 1490 AGAAAGCTTGGACACCTACTGTTTCTAGTCTTGTGCTTGTGCTTGTGCTTGTGCTA 1549  
Db 541 AGAAAGCTTGGACACCTACTGTTTCTAGTCTTGTGCTTGTGCTTGTGCTTGTGCTA 600  
Qy 1550 TCTACCTGTATTTTGTGCTGCTGAGACCAAGAAACAGAACTATGCAAGAAATCAGCCAGG 1609  
Db 601 CCTACTTCTATTTGTCTCTCCGAGACCAAGAAACAGAACTATGCAAGAAATCAGCCAGG 659  
Qy 1610 CATTTTCCAAAGGAAACAAAGCTATCCACAGAGAGAA 1649  
Db 660 CATTTTCCAAAGGAAACAAAGCTATCCACAGAGAGAA 699

RESULT 13  
BE740763  
LOCUS 601593559F1 NIH\_MGC\_9 Homo sapiens cDNA clone IMAGE:3947511 5',  
DEFINITION BE740763 893 bp mRNA linear EST 15-SEP-2000  
mRNA sequence.  
ACCESSION BE740763  
VERSION BE740763.1 GI:10154755  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS 1 (bases 1 to 893)  
TITLE NIH-MGC http://mgi.nci.nih.gov/  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: DCTD/DTF  
cDNA Library Preparation: Ling Hong/Rubin Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov  
Plate: LLCM809 row: e column: 16  
High quality sequence stop: 734.  
Location/Qualifiers  
1..893  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3947511"  
/tissue\_type="adenocarcinoma cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_9"  
/note="Organ: ovary; Vector: pOTB7; Site 1: XhoI; Site 2:  
EcoRI; cDNA made by oligo-dT priming. Directionally  
cloned into EcoRI/XhoI sites using the following 5'  
adaptor: GGCACGAG(G). Size-selected >500bp for average  
insert size 1.8kb. Library constructed by Ling Hong in  
the laboratory of Gerald M. Rubin (University of  
California, Berkeley) using ZAP-cDNA synthesis kit  
(Stratagene) and Superscript II RT (Life Technologies)."  
ORIGIN  
Query Match 20.1%; Score 471.6; DB 10; Length 893;  
Best Local Similarity 90.6%; Pred. No. 1.5e-55;  
Matches 560; Conservative 0; Mismatches 49; Indels 9; Gaps 5;

Qy 217 AGAAAGGACTGCTCTGCTGCTCTGCTGCTCTCGCGGCGCCTTGGCTCTCC 276  
Db 202 AAATTGGACTGCTCTGCTGCTCTGCTGCTCTCGCGGCGCCTTGGCTCTCC 261  
Qy 277 TTCTCTACGGCTACAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 336  
Db 262 TTCTCTACGGCTACAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTT 321  
Qy 337 TACATGAGTCAATGGAAGAAGGATGAGCGTCCCAATAGACCCAGACACTCTG 396  
Db 322 TACATGAGTCAATGGAAGAAGGATGAGCGTCCCAATAGACCCAGACACTCTG 381  
Qy 397 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 456  
Db 382 CTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441  
Qy 457 AAGATGATTGGAAGGTTCTTGGGAGGAGACACATTTGCTGGCCCAATAATGG 516  
Db 442 AAGATGATTGGAAGGTTCTTGGGAGGAGACACATTTGCTGGCCCAATAATGG 501  
Qy 517 ATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 576  
Db 502 ATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561  
Qy 577 GTGGAGCGCTTCAATCATGAGGATAGATGAGGCGCTGCTGCTGCTGCTGCTG 636  
Db 562 GTGGAGCGCTTCAATCATGAGGATAGATGAGGCGCTGCTGCTGCTGCTGCTG 620  
Qy 637 CTCAGTGAGATCTCACCCCAAGG--AGATCCGCTGCTGCTGCTGCTGCTGCTG 693  
Db 621 CTCAGTGAGATCTCACCCCAAGGATCCCGTGGCTCTCTGCTGCTGCTGCTGCT 680  
Qy 694 TTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750  
Db 681 TTTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 740  
Qy 751 AAGGAGAGTACCTGGCCATACCTGTTGAGTATGTTGCTGCTGCTGCTGCTGCT 810  
Db 741 GACAGGGAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 798  
Qy 811 CTGAGCGCTTCCCTTCTC 828  
Db 799 GTGGGCGCTTCCCTTCTC 816

RESULT 14

Db	554	G-CAGGTGGCATCCCGTTTCATCTGGACTTGGTGGTTCCTCCACGCATCTCAGCGGGG	711
<p>BE741795 1017 bp mRNA linear EST 15-SEP-2000</p> <p>601595653F1 NIH_MGC_9 Homo sapiens cDNA clone IMAGE:3949549 5', mRNA sequence.</p> <p>BE741795</p> <p>BE741795.1 GI:10155787</p> <p>EST.</p> <p>Homo sapiens (human)</p> <p>Homo sapiens</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.</p> <p>NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a></p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished (1999)</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a></p> <p>Tissue Procurement: DCTD/DTF</p> <p>cDNA Library Preparation: Ling Hong/Rubin Laboratory</p> <p>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Incyte Genomics, Inc.</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">image.llnl.gov</a></p> <p>Plate: L10C814 row: j column: 14</p> <p>High quality sequence start: 34</p> <p>High quality sequence stop: 828.</p>			
<p>FEATURES</p> <p>source</p> <p>1. 1017</p> <p>/organism="Homo sapiens"</p> <p>/mol_type="mRNA"</p> <p>/db_xref="taxon:9606"</p> <p>/clone="IMAGE:3949549"</p> <p>/tissue_type="adenocarcinoma cell line"</p> <p>/lab_host="DH10B (phage-resistant)"</p> <p>/clone_lib="NIH_MGC_9"</p> <p>/note="Organ: ovary; Vector: pOT7; Site 1: XhoI; Site 2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCGAGGAG(G). Size-selected &gt;500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."</p>			
<p>ORIGIN</p> <p>Query Match 24.9%; Score 584; DB 10; Length 1017;</p> <p>Best Local Similarity 98.1%; Pred. No. 5.2e-71;</p> <p>Matches 612; Conservative 0; Mismatches 10; Indels 2; Gaps 2;</p>			
Qy	217	AGAAAGGACTGGTCTGCTCGTCTCGTGGCTCCCTCGGGGGCGCTTCGGCTCTCC	276
Db	252	AAATGGGACTGGTCTGCTCGTCTCGTGGCTCCCTCGGGGGCGCTTCGGCTCTCC	311
Qy	277	TTCTCTACGGCTACAACCTGTCGGTGGTGAATGCCGCCACCCGCTACATCAAGGCCCTT	336
Db	312	TTCTCTACGGCTACAACCTGTCGGTGGTGAATGCCGCCACCCGCTACATCAAGGCCCTT	371
Qy	337	TACAATGAGTCATGGGAAGAAGGATGGAGTCCATAGACCAGACACTCTGACTCTG	396
Db	372	TACAATGAGTCATGGGAAGAAGGATGGAGTCCATAGACCAGACACTCTGACTCTG	431
Qy	397	CTCTGGTCTGACTGTCCTCCATATTCGATCCGATGGTGGACGTTAATTGTG	456
Db	432	CTCTGGTCTGACTGTCCTCCATATTCGATCCGATGGTGGACGTTAATTGTG	491
Qy	457	AAATGATTGGAAGGTTCTTGGGAGGAAGACACATTTGCTGGCCAAATATGGGTTTGA	516
Db	492	AAATGATTGGAAGGTTCTTGGGAGGAAGACACATTTGCTGGCCAAATATGGGTTTGA	551
Qy	517	ATTTCTGCTGATTCGCTGATGGCTTCGCTCCAGGCGAGAGCCTTTGAAATGCTCAT	576
Db	552	ATTTCTGCTGATTCGCTGATGGCTTCGCTCCAGGCGAGAGCCTTTGAAATGCTCAT	611
Qy	577	GTGGGACGCTTCATCATGGGCATAGATGGAGCGTCGCCCTCAGTGTGCTCCCATGTAC	636
Db	612	GTGGGACGCTTCATCATGGGCATAGATGGAGCGTCGCCCTCAGTGTGCTCCCATGTAC	671
Qy	637	CTCAGTGAGATCTACCCCAAGAGATCCGTGGCTCTCTGGGGCAGGTGAC-TGCCATCTT	695
Db	672	CTCAGTGAGATCTACCCCAAGAGATCCGTGGCTCTCTGGGGCAGGTGAC-TGCCATCTT	731
Qy	696	TATCTGCAATGGCGTGTTCACCTGGGCGAGCTTCTGGGCGTCCCGAGCTGTGGGAAA-GG	754
Db	732	TATCTGCAATGGCGTGTTCACCTGGGCGAGCTTCTGGGCGTCCCGAGCTGTGGGAAA-GG	791
Qy	755	AGAGTACCTGGCCATACCTGTTTGGAGTGAATGTGGTCCCTGCCGCTTCCAGCTGTGTA	814
Db	792	AGAGTACCTGGCCATACCTGTTTGGAGTGAATGTGGTCCCTGCCGCTTCCAGCTGTGTA	851
Qy	815	GCCTTCCTCTCTCCCGACAGCC 838	
Db	852	AGCTCCCTCTCTCCCGACAGCC 875	
<p>RESULT 12</p> <p>BI154904</p> <p>LOCUS</p> <p>DEFINITION</p> <p>602902896F1 NCI_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5032582 5', mRNA sequence.</p> <p>BI154904</p> <p>BI154904.1 GI:14614905</p> <p>EST.</p> <p>Mus musculus (house mouse)</p> <p>Mus musculus</p> <p>Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.</p> <p>NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a></p> <p>National Institutes of Health, Mammalian Gene Collection (MGC)</p> <p>Unpublished (1999)</p> <p>Contact: Robert Strausberg, Ph.D.</p> <p>Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a></p> <p>Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.</p> <p>cDNA Library Preparation: Life Technologies, Inc.</p> <p>cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)</p> <p>DNA Sequencing by: Incyte Genomics, Inc.</p> <p>Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a></p> <p>Plate: LLAM1090 row: p column: 23</p> <p>High quality sequence stop: 683.</p>			
<p>FEATURES</p> <p>source</p> <p>1. 801</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/strain="129, C57BL/6J, FVB/N"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="IMAGE:5032582"</p> <p>/tissue_type="tumor, gross tissue"</p> <p>/dev_stage="10 months"</p> <p>/lab_host="DH103"</p> <p>/clone_lib="NCI_CGAP_Mam3"</p> <p>/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies. Investigators providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."</p>			
<p>ORIGIN</p> <p>Query Match 21.7%; Score 508.8; DB 12; Length 801;</p> <p>Best Local Similarity 83.9%; Pred. No. 1.3e-60;</p>			











Db 460 ATCTTGACTGGTGGTTCCTTCCAGCAATCTCAGCGGCGGCTGCTTCATCATTCATTCGAGGC 401

Qy 1438 ACCGTCAACTGGCTCTCCAACTTTGCTGTGGGCTCTCTTCCATTCATTCAGAAAGT 1497

Db 400 ACCGTCAACTGGCTCTCCAACTTTGCTGTGGGCTCTCTTCCATTCATTCAGAAAGT 341

Qy 1498 CTGGACACCTACTGTTTCTTCTAGTCTTTGCTACAAATTTGTATCAGAGTGTCTATCTACTG 1557

Db 340 CTGGACACCTACTGTTTCTTCTAGTCTTTGCTACAAATTTGTATCAGAGTGTCTATCTACTG 281

Qy 1558 TATTTTGTCTGCTGCTGAGACCAAAACAGAACCTTATCGAGAAATCAGCAGGCAATTTTCC 1617

Db 280 TATTTTGTCTGCTGCTGAGACCAAAACAGAACCTTATCGAGAAATCAGCAGGCAATTTTCC 221

Qy 1618 AAAAGGAACAAAGCATATCCACCACAGAGAGAGAAATCGACTCAGCTGTCTACTGTATGCT 1674

Db 220 AAAGGACAAAGCATATCCACCACAGAGAGAGAAATCGACTCAGCTGTCTACTGTATGCT 164

RESULT 6

LOCUS BI760988

DEFINITION 603048508P1 NIH\_MGC\_116 Homo sapiens cDNA clone IMAGE:5188849 5',

ACCESSION BI760988

VERSION BI760988.1 GI:15752566

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC http://mgi.nci.nih.gov/

JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)

COMMENT Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cga@b6-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

Clone sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov

Plate: LIA411471 row: p column: 02

High quality sequence start: 5

High quality sequence stop: 859.

FEATURES

Location/Qualifiers

1..876

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5188849"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_116"

/note="Organ: Pooled colon, kidney, stomach; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 26 yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 32.3%; Score 757.2; DB 12; Length 876;

Best Local Similarity 96.0%; Pred. No. 8.8e-95;

Matches 831; Conservative 0; Mismatches 28; Indels 7; Gaps 5;

Qy 79 AGGAACAAATAGGAATTCCAA-GGAACCTGGGCTAGT-TCCCTCACAGATGACACCA 136

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)

COMMENT On Feb 16, 2001 this sequence version replaced gi:12929891.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

Library was constructed by Life Technologies, a division of

Invitrogen. This sequence belongs to sequence cluster 2476.r For

more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DI026CD12NP1&cluster=2476.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/

Invitrogen Corporation 1600

Paraday Avenue Genoscope sequence ID : CS0DI026CD12NP1.

FEATURES

Location/Qualifiers

1..940

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CS0DI026YH23"

/tissue\_type="PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN

Query Match 33.0%; Score 773.4; DB 9; Length 940;

Best Local Similarity 99.6%; Pred. No. 5.1e-97;

Matches 774; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 898 TTCTTGGTAAAGCAGACGTTTCCCAAGAGGTAGAGAGTCTCTGCTGAGACCAAGTGG 957

Db 940 TTCTTGGTAAAGCAGACGTTTCCCAAGAGGTAGAGAGTCTCTGCTGAGACCAAGTGG 881

Qy 958 CAGAGAGCATCCGCTGTGTCGGTGTGGAGTCTCTGAGAGTCCCTACGTCGCTGG 1017

Db 880 CAGAGAGCATCCGCTGTGTCGGTGTGGAGTCTCTGAGAGTCCCTACGTCGCTGG 821

Qy 1018 CAGGTGTCACCTGATTTGTCACATGGCTGCTACCATGCTGTGGCTCTCAATGCAAT 1077

Db 820 CAGGTGTCACCTGATTTGTCACATGGCTGCTACCATGCTGTGGCTCTCAATGCAAT 761

Qy 1078 TGCTTCTATACCAACAGCATCTTTGGAAGGTGGGATCCCTCCGCAAGATGCCATAC 1137

Db 760 TGCTTCTATACCAACAGCATCTTTGGAAGGTGGGATCCCTCCGCAAGATGCCATAC 701

Qy 1138 GTCACCTTGAGTACAGGGGATCGAGATTTGGCTGCGCTCTCTCTGTTGGTCAAT 1197

Db 700 GTCACCTTGAGTACAGGGGATCGAGATTTGGCTGCGCTCTCTCTGTTGGTCAAT 641

Qy 1198 GAGCAGCTGGGACGAGACCCCTCTCATGTTGGCTTTGGGCTCATGGGCTCTCTTT 1257

Db 640 GAGCAGCTGGGACGAGACCCCTCTCATGTTGGCTTTGGGCTCATGGGCTCTCTTT 581

Qy 1258 GGGACCTCACCATCAGCTGACCTTCGAGGACCGCCCTGGGTCCCTACCTGAGT 1317

Db 580 GGGACCTCACCATCAGCTGACCTTCGAGGACCGCCCTGGGTCCCTACCTGAGT 521

Qy 1318 ATCGTGGGCAATCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTC 1377

Db 520 ATCGTGGGCAATCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTC 461

Qy 1378 ATCTTGAGTGGTGGTCTTTCCAGCAATCTCAGCGGCGGCTGCTTCATCATTTGAGGC 1437



[illegible]

Tissue Procurement: Life Technologies, Inc.			
CDNA Library Preparation: Life Technologies, Inc.			
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)			
DNA Sequencing by: Incyte Genomics, Inc.			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILLNL at: <a href="http://image.llnl.gov">http://image.llnl.gov</a>			
Plate: LILM11455 row: k column: 16			
High quality sequence stop: 848.			
Location/Qualifiers			
source	1.. 968		
	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5182599"		
	/lab_host="DH10B"		
	/clone_lib="NIH_MGC_115"		
	/note="Organ: pooled colon, kidney, stomach; Vector: pCMV-SPORE6; Site:1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 3 colons, age 2s yo male, 49 yo female, 71 yo male colon; 46 yo male kidney, and pool of 2 stomachs, 62 yo male and 70 yo female. Library is oligo-dT primed and directionally cloned [EcoRV site is destroyed upon cloning]. Average insert size 1.4 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 023. Note: this is a NIH_MGC Library."		
ORIGIN			
	Query Match	35.5%;	Score 832; DB 12; Length 968;
	Best Local Similarity	96.5%;	Pred. No. 4,4e-105;
	Matches	915;	Conservative 0; Mismatches 25; Indels 8; Gaps 6;
Qy	39	CTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGGCAAGAAACAAATAGGAATTC	98
Db	1	CTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGGCAAGAAACAAATAGGAATTC	60
Qy	99	CA-AGGAACCTGGCCCTAGTTCCCTCACATGACACACCGCCGCGGCCCTCCAGGGC	157
Db	61	CAGAGMACTGGCCCTAGTTCCCTCACATGACACACCGCCGCGGCCCTCCAGGGC	120
Qy	158	CAGGAGGGGCATCTCTGGAGTGTGACACCTGAGGAGTGGGGTCCAGGTGGAGGAGAA	217
Db	121	CAGGAGGGGCATCTCTGGAGTGTGACACCTGAGGAGTGGGGTCCAGGTGGAGGAGAA	180
Qy	218	GAAAGACTGGTCTCTGCTCGTCTCTGCTGTCCTCCCTCGCGGGGGCTTCGGCTCCTCCT	277
Db	181	GAAAGACTGGTCTCTGCTCGTCTCTGCTGTCCTCCCTCGCGGGGGCTTCGGCTCCTCCT	240
Qy	278	TCCTCTACGGCTACAACTGTTCGGTGGTGAATGCCGCCACCCCGTATACATCAAGGCCTTTT	337
Db	241	TCCTCTACGGCTACAACTGTTCGGTGGTGAATGCCGCCACCCCGTATACATCAAGGCCTTTT	300
Qy	338	ACAAATGAGTATCGGAAAGAGGCAATGAGTCCAAATAGACCCAGACACTCTGACTCTGTC	397

338	QY	ACAATGAGTATGGGAAAGAGCATGGACGTCUCAAATAGACCCGAGACACATCTGACATCTGC	397
301	Db	ACAATGAGTCATGGGAAAGAGCATGGACGTCCTCAATAGACCCGAGACACATCTGACATCTGC	360
398	QY	CTCGGTCTGTGACTGTGTCCATATTCGCCATCGGTGGACCTTGTGGGAGCGTTAAATTGTGA	457
361	Db	CTCGGTCTGTGACTGTGTCCATATTCGCCATCGGTGGACCTTGTGGGAGCGTTAAATTGTGA	420
458	QY	AGATGATTCGGAAAGGTCCTTGGGAGAGACACACTTTGCTGSCCATATATGGGTTTGCAC	517
421	Db	AGATGATTCGGAAAGGTCCTTGGGAGAGACACACTTTGCTGSCCAATATATGGGTTTGCAC	480
518	QY	TTTCTCGCTGATGTGATGGCGCTGCTCGCTCCAGCAGAGAGCGCTTTGAAAAGCTCATTTG	577
481	Db	TTTCTCGCTGATGTGATGGCGCTGCTCGCTCCAGCAGAGAGCGCTTTGAAAAGCTCATTTG	540
578	QY	TGGGAGCGCTTCATCATGGGCAATAGATGGAGGGGCTGCCCTTCAGTGTGCTCCCCCATGTACC	637
541	Db	TGGGAGCGCTTCATCATGGGCAATAGATGGAGGGGCTGCCCTTCAGTGTGCTCCCCCATGTACC	600

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
Faraday Avenue Genoscope sequence ID : CS0DI026CD12QPI.  
Location/Qualifiers  
1. 932  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DI026YH23"  
/tissue type="PLACENTA COT 25-NORMALIZED"  
/clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo (dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and cloned into the Not I and EcoR V  
sites of the pCMVSPORT 6 vector. Library was normalized."

FEATURES  
source  
Query Match 36.4%; Score 852.2; DB 9; Length 932;  
Best Local Similarity 99.6%; Pred. No. 7.6e-108;  
Matches 854; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

ORIGIN  
QY 17 GCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATGAGTGGTCACTGAGACCCATGG 76  
DB 76 GCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATGAGTGGTCACTGAGACCCATGG 135  
QY 77 CAAGGAAACAAATAGGAATCCAGGAATCTGGGCTAGTTCCTCTACAGATCACACCA 136  
DB 136 CAAGGAAACAAATAGGAATCCAGGAATCTGGGCTAGTTCCTCTACAGATCACACCA 195  
QY 137 GCCACGCCGGCTCCAGGGCCAGGGAGGACACTGCTGAGTGTGACCACTGAGGAGTG 196  
DB 196 GCCACGCCGGCTCCAGGGCCAGGGAGGACACTGCTGAGTGTGACCACTGAGGAGTG 255  
QY 197 GGGTCCAGGTGGAAGAGGAGAAAGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 256  
DB 256 GGGTCCAGGTGGAAGAGGAGAAAGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 315  
QY 257 GGGGGCCCTGGCTCCCTCTCTACGCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 316  
DB 316 GGGGGCCCTGGCTCCCTCTCTACGCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTG 375  
QY 317 CCCCCTACATCAAGCCCTTTTACATGAGTCAATGGAAAGAGGAGGAGGAGGAGGAGGAG 376  
DB 376 CCCCCTACATCAAGCCCTTTTACATGAGTCAATGGAAAGAGGAGGAGGAGGAGGAGGAG 435  
QY 377 ACCCAGACACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 436  
DB 436 ACCCAGACACTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 495  
QY 437 TTGTGGGACGTTAATCTGAGATGATTGAAAGGTTCTTGGAGGAGGAGGAGGAGGAGG 496  
DB 496 TTGTGGGACGTTAATCTGAGATGATTGAAAGGTTCTTGGAGGAGGAGGAGGAGGAGG 555  
QY 497 TGGCAATATGGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556  
DB 556 TGGCAATATGGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 615  
QY 557 GAGCCTTTGAAATGCTCATTTGGGAGCGCTTCATCATGGGATAGATGGAGGCGTGCACC 616  
DB 616 GAGCCTTTGAAATGCTCATTTGGGAGCGCTTCATCATGGGATAGATGGAGGCGTGCACC 675  
QY 617 TCAAGTGTCTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 676  
DB 676 TCAAGTGTCTCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 735  
QY 677 GGCAGGTGACATGCTATTTATCTGCAATGCGGTGTTCACTGGGAGGAGGAGGAGGAG 736  
DB 736 GGCAGGTGACATGCTATTTATCTGCAATGCGGTGTTCACTGGGAGGAGGAGGAGGAG 795  
QY 797 CCGAGCTCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTGAATGCTGCTGCTG 796  
DB 796 CCGAGCTCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTGAATGCTGCTGCTG 855  
QY 797 CCGTGTCCAGTGTGAGCCCTTCCCTTTCTCCGGGACAGCCACGCTACCTGCTCTTGG 856

Db 856 CCCTTGTCCAGTCTGAGCCTTCCCTTTCCTCCGACACGCCACGCTACCTGCTCTGG 915  
QY 857 AGAAGCACACGAGGCA 873  
DB 916 AGAAGCACACGAGGCA 932

RESULT 2  
BM545063 1131 bp mRNA linear EST 20-FEB-2002  
LOCUS AGENCOURT\_6497736 NIH\_MGC\_125 Homo sapiens cDNA clone IMAGE:558878  
DEFINITION 5', mRNA sequence.  
ACCESSION BM545063  
VERSION BM545063.1 GI:18776860  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 1131)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs@mail.nih.gov  
Tissue Procurement: Invitrogen  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM12360 row: k column: 23  
High quality sequence start: 28  
High quality sequence stop: 725.

FEATURES  
Location/Qualifiers  
1. 1131  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:558878"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_125"  
/note="Organ: ovary (pool of 3); Vector: pCMV-SPORT6;  
Site 1: EcoRV (destroyed); Site 2: NotI; RNA source pool  
of three ovaries, from females ranging in age from 38 to  
49 yo. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 2.1 kb, insert size range 1-3.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 036."

ORIGIN  
Query Match 36.2%; Score 848.8; DB 12; Length 1131;  
Best Local Similarity 96.0%; Pred. No. 2e-107;  
Matches 968; Conservative 0; Mismatches 27; Indels 13; Gaps 9;

QY 313 CCCACCCCTACATCAAGCCCTTTTACATGAGTCAATGGAAAGAGGAGGAGGAGGAGGAGG 372  
DB 28 CCCACCCCTACTTC-AGGCCCTTTTACATGAGTCAATGGAAAGAGGAGGAGGAGGAGG 86  
QY 373 ATAGACCCAGACACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432  
DB 87 ATAGACCCAGACACTGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 146  
QY 433 GGACTTGTGGGACCGTTAATTTGTGAAGATGATTGGAAGGTTCTTGGGAGGAGGAGGAGG 492  
DB 147 GGACTTGTGGGACCGTTAATTTGTGAAGATGATTGGAAGGTTCTTGGGAGGAGGAGG 206  
QY 493 TTGCTGGCCAAATATGGGTTTCCAAATTTCTGCTGATGCTGCTGCTGCTGCTGCTGCTG 552

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 22:38:47 ; Search time 5970 Seconds  
(without alignments)  
11719.780 Million cell updates/sec

Title: US-09-981-947B-1  
Perfect score: 2343  
Sequence: 1 tcgaccacgcgtccgctct.....aaaaaaaaagggcgccgc 2343

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_ges\_hum:\*  
18: em\_ges\_inv:\*  
19: em\_ges\_pln:\*  
20: em\_ges\_vrt:\*  
21: em\_ges\_fun:\*  
22: em\_ges\_mam:\*  
23: em\_ges\_mus:\*  
24: em\_ges\_pro:\*  
25: em\_ges\_rod:\*  
26: em\_ges\_phg:\*  
27: em\_ges\_vrl:\*  
28: gb\_ges1:\*  
29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	852.2	36.4	932	9	AL546827
2	848.8	36.2	1131	12	BM545063
3	832	35.5	968	12	BI517809
4	778.2	33.2	929	12	BI911320

C	5	773.4	33.0	940	9	AL572022
	6	757.2	32.3	876	12	BI760988
	7	754.2	32.2	889	12	BI821650
	8	752.8	32.1	2901	11	AK037093
C	9	669.2	28.6	687	14	CA309954
	10	588.6	25.1	916	10	BF035577
	11	584	24.9	1017	10	BE741795
	12	508.8	21.7	801	12	BI154904
	13	471.6	20.1	893	10	BE740763
	14	452.8	19.3	1297	12	BM925071
	15	439.4	18.8	961	10	BE745805
	16	432.2	18.4	2473	11	AK033725
C	17	428.4	18.3	675	14	CA307279
	18	427.6	18.3	1539	29	AY401415
	19	424.4	18.1	3097	11	AK029720
C	20	422.4	18.0	690	12	BM678584
	21	419.6	17.9	2100	11	AF289610
	22	415.2	17.7	667	10	BE625357
	23	411.2	17.6	597	10	BE305606
	24	408.6	17.4	682	12	BG971798
C	25	408.4	17.4	620	12	BQ015082
C	26	407.4	17.4	595	10	AW300078
	27	403.6	17.2	602	9	AI766064
C	28	397.4	17.0	629	10	BB660933
C	29	391.6	16.7	605	10	BE463553
C	30	387.8	16.6	597	10	AW263084
	31	387.8	16.6	1542	29	AY401417
	32	378.8	16.2	635	10	BB660981
C	33	376.8	16.1	962	12	BI762013
C	34	375.4	16.0	571	9	AI671682
C	35	368.8	15.7	641	10	BB664247
C	36	365.2	15.6	608	10	BE464196
C	37	356.2	15.2	563	9	AI433678
C	38	355.2	15.2	1506	29	AY410442
C	39	316.4	13.5	495	9	AI312682
C	40	313.2	13.4	602	10	BB630190
	41	311.6	13.3	770	14	CB521349
	42	306.6	13.1	339	14	FO0548
C	43	304.4	13.0	498	10	AW614815
C	44	302.8	12.9	500	9	AI022719
C	45	300.4	12.8	1453	29	AY401416

ALIGNMENTS

RESULT 1  
AL546827  
LOCUS  
DEFINITION AL546827 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA clone CSODI026VH23 5-PRIME, mRNA sequence.  
ACCESSION AL546827  
VERSION AL546827.2 GI:31268660  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
1 (bases 1 to 932)  
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:12880321.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 EVRY cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 2476.r For more information about this cluster, see  
http://www.genoscope.cns.fr/  
cgi-bin/cluster.cgi?seq=CSODI026CDI2QPI&cluster=2476.r. Contact : Feng Liang Email : fliang@lifetech.com URL :

Search completed: April 8, 2004, 22:43:36  
Job time : 839 secs



QY 1261 ACCCTACCAATCAGCTGACCTGACGACCAACGCCCCCTGGTCCCTACCTGAGTATC 1320  
Db 812 ACCCTACCAATCAGCTGACCTGACGACCAACGCCCCCTGGTCCCTACCTGAGTATC 871  
QY 1321 GTGGGCAATCTGGCCATCAGCTCTTTCTGAGTGGGCGGCGAGTGGGATCCCGTTATC 1380  
Db 872 GTGGGCAATCTGGCCATCAGCTCTTTCTGAGTGGGCGGCGAGTGGGATCCCGTTATC 931  
QY 1381 TTGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCAATCAATTCAGGCACC 1440  
Db 932 TTGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCAATCAATTCAGGCACC 991  
QY 1441 GTCACTGGCTTCCAACTTTGTGTGGGCTCTTTCCCAATCAATTCAGAAAAAGTCTG 1500  
Db 992 GTCACTGGCTTCCAACTTTGTGTGGGCTCTTTCCCAATCAATTCAGAAAAAGTCTG 1051  
QY 1501 GACACCTACTGTTCTTCTAGCTTTTGTCTACAAATTTGTATCAGAGTGGCTATCTACCTGAT 1560  
Db 1052 GACACCTACTGTTCTTCTAGCTTTTGTCTACAAATTTGTATCAGAGTGGCTATCTACCTGAT 1111  
QY 1561 TTTGTGCTGCTGAGACCAAAACAGAACTATGACAGAAATCAGCCAGGCAATTTTCCAAA 1620  
Db 1112 TTTGTGCTGCTGAGACCAAAACAGAACTATGACAGAAATCAGCCAGGCAATTTTCTCAG 1171  
QY 1621 AGGAA 1625  
Db 1172 GGCAA 1176

RESULT 13

US-10-188-186-87  
; Sequence 87, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 87  
; LENGTH: 1189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)..(1189)  
US-10-188-186-87

Query Match 48.3%; Score 1130.6; DB 12; Length 1189;

Best Local Similarity 99.2%; Pred. No. 9.2e-301;  
Matches 1136; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
QY 481 AGGAAGCAGACATTTGCTGGCCCAATTAATGGGTTTGCATTTCTGCTGCATTTGCTGATGACC 540  
Db 32 AGGAAGCAGACATTTGCTGGCCCAATTAATGGGTTTGCATTTCTGCTGCATTTGCTGATGACC 91  
QY 541 TGCTCGCTCCAGCAGGAGCCCTTTGAAATGCTCATTTGTGGGACGCTTCAATGAGGCATA 600  
Db 92 TGCTCGCTCCAGCAGGAGCCCTTTGAAATGCTCATTTGTGGGACGCTTCAATGAGGCATA 151  
QY 601 GATGAGGCGTGGCCCTCAGTGTCTGCCCAATGATACCTGATGAGATCTCAACCAAGAG 660  
Db 152 GATGAGGCGTGGCCCTCAGTGTCTGCCCAATGATACCTGATGAGATCTCAACCAAGAG 211  
QY 661 ATCCGTTGGCTCTCTGGGCGAGGTGACTGCCATCTTTATCTGCATTTGGGCTGTTCACTGG 720  
Db 212 ATCCGTTGGCTCTCTGGGCGAGGTGACTGCCATCTTTATCTGCATTTGGGCTGTTCACTGG 271  
QY 721 CAGCTTCTGGGCTGCGCCAGTGTCTGGGAAAGAGAGTACTGTGCCCATACCTGTTTGA 780  
Db 272 CAGCTTCTGGGCTGCGCCAGTGTCTGGGAAAGAGAGTACTGTGCCCATACCTGTTTGA 331  
QY 781 GTGATTTGGTCCCTGGCGTGTCTCAGTGTGTGAGCTTCCCTTCTCCGAGACAGGCCA 840  
Db 332 GTGATTTGGTCCCTGGCGTGTCTCAGTGTGTGAGCTTCCCTTCTCCGAGACAGGCCA 391  
QY 841 CGCTACCTGCTCTTGGAGAAAGCACAACGAGGCAAGAGCTGTGAAGCTTCCCAACGTT 900  
Db 392 CGCTACCTGCTCTTGGAGAAAGCACAACGAGGCAAGAGCTGTGAAGCTTCCCAACGTT 451  
QY 901 TTGGTAAAGCAGACGTTTCCCAAGAGTGAAGAGTCTCTGGGTGAGAGCCACGTCAG 960  
Db 452 TTGGTAAAGCAGACGTTTCCCAAGAGTGAAGAGTCTCTGGGTGAGAGCCACGTCAG 511  
QY 961 AGGAGCATCGCTGGTGTCTGGAGTGTCTGAGCTGTCTGAGAGCTCCCTAGCTGGTGGCAG 1020  
Db 512 AGGAGCATCGCTGGTGTCTGGAGTGTCTGAGAGTGTCTGAGAGTCCCTAGCTGGTGGCAG 571  
QY 1021 GTGGTCACCGTCAATTTGTCACCAATGGCTGTACAGCTCTCTGGGCTCAATGCAATTTGG 1080  
Db 572 GTGGTCACCGTCAATTTGTCACCAATGGCTGTACAGCTCTCTGGGCTCAATGCAATTTGG 631  
QY 1081 TTCTATACCAACAGCATCTTTTGGAAAGCTGGGATCCCTCCGCAAGAGATCCCAATGTC 1140  
Db 632 TTCTATACCAACAGCATCTTTTGGAAAGCTGGGATCCCTCCGCAAGAGATCCCAATGTC 691  
QY 1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGGCGTCTCTCTGTTGGTGGTCAATGAG 1200  
Db 692 ACCTTGAGTACAGGGGGCATCGAGACTTTGGCTGGCGTCTCTCTGTTGGTGGTCAATGAG 751  
QY 1201 CACCTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTTGGG 1260  
Db 752 CACCTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTTGGG 811  
QY 1261 ACCCTCACCATCAAGCTGACCTTGAGGACCAACGCCCCCTGGGTCCCTACCTGAGTATC 1320  
Db 812 ACCCTCACCATCAAGCTGACCTTGAGGACCAACGCCCCCTGGGTCCCTACCTGAGTATC 871  
QY 1321 GTGGGATTTGGCCATCATGCTCTTTCTGAGTGGGCGGAGGTGGCATCCCTTCATC 1380  
Db 872 GTGGGATTTGGCCATCATGCTCTTTCTGAGTGGGCGGAGGTGGCATCCCTTCATC 931  
QY 1381 TTGACTGGTGGTCTTCTCCAGCAATCTCAGCGCGGCGTGGCTTTCATCATTCAGGAGCAC 1440  
Db 932 TTGACTGGTGGTCTTCTCCAGCAATCTCAGCGCGGCGTGGCTTTCATCATTCAGGAGCAC 991  
QY 1441 GTCAACTGGCTTCCAACTTTGTGTGGGCTCTCTTCCCAATCAATTCAGAAAAAGTCTG 1500  
Db 992 GTCAACTGGCTTCCAACTTTGTGTGGGCTCTCTTCCCAATCAATTCAGAAAAAGTCTG 1051  
QY 1501 GACACCTACTGTTCTTCTAGCTTTTGTCTACAAATTTGTATCAGAGTGGCTATCTACCTGAT 1560



961 AGGAGCATCGGCTGGTGTGCTGGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAG 1020  
512 AGGAGCATCGGCTGGTGTGCTGGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAG 571  
1021 GTGGTCACCGTGAATGTACCAATGGCTGCTACCAAGCTCTGTGGGCTCAATGCAATTTGG 1080  
572 GTGGTCACCGTGAATGTACCAATGGCTGCTACCAAGCTCTGTGGGCTCAATGCAATTTGG 631  
1081 TTCTATACCAAGAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 1140  
632 TTCTATACCAAGAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 691  
1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGGTGCGCTCTTCTCTGGTTTGGTCAATTGAG 1200  
692 ACCTTGAGTACAGGGGGCATCGAGACTTTGGGTGCGCTCTTCTCTGGTTTGGTCAATTGAG 751  
1201 CACCTGGAGCGGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 1260  
752 CACCTGGAGCGGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 811  
1261 ACCTCACCATCAGCGTGAACCTTGAGACCAAGCGCCCTGGGTCCCTACCTGAGTATC 1320  
812 ACCTCACCATCAGCGTGAACCTTGAGACCAAGCGCCCTGGGTCCCTACCTGAGTATC 871  
1321 GTGGGATTTCTGGCCATCATGCGCTCTTCTGAGTGGCCAGGTGGCATCCCGTTTCATC 1380  
872 GTGGGATTTCTGGCCATCATGCGCTCTTCTGAGTGGCCAGGTGGCATCCCGTTTCATC 931  
1381 TTGACTGTGTGAGTTCTTCAGCAATCTCAGCGCGCGCTGCTTCATCATTCAGGACCC 1440  
932 TTGACTGTGTGAGTTCTTCAGCAATCTCAGCGCGCGCTGCTTCATCATTCAGGACCC 991  
1441 GTCAACTGGCTCTCCAACTTTCTGTGGGCTCTCTTCCCATCATTCAGAAAGTCTG 1500  
992 GTCAACTGGCTCTCCAACTTTCTGTGGGCTCTCTTCCCATCATTCAGAAAGTCTG 1051  
1501 GACACCTACTGTTTCTAGTCTTTGTGTAGCAATTTGTATACAGGTGCTATCTACCTGAT 1560  
1052 GACACCTACTGTTTCTAGTCTTTGTGTAGCAATTTGTATACAGGTGCTATCTACCTGAT 1111  
1561 TTGTGCTCCCTGAGACCAAAACAGAACTATGAGAAATCAGACGAGCATTTTCCAAA 1620  
1112 TTGTGCTCCCTGAGACCAAAACAGAACTATGAGAAATCAGACGAGCATTTTCCGAG 1171  
1621 AGGAA 1625  
1172 GGCAA 1176

RESULT 12  
US-10-188-186-83  
; Sequence 83, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11

; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 83  
; LENGTH: 1189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)...(1189)  
US-10-188-186-83

Query Match 48.3%; Score 1130.6; DB 12; Length 1189;  
Best Local Similarity 99.2%; Pred. No. 9.2e-301;  
Matches 1136; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
Qy 481 AGGAAGCACACTTTGGTGGCCAAATAATGGTGTTCGCAATTTCTGCTGCTGATGGCC 540  
Db 32 AGGAAGCACACTTTGGTGGCCAAATAATGGTGTTCGCAATTTCTGCTGCTGATGGCC 91  
Qy 541 TGCTCGCTCCAGGAGAGAGCCTTTGAAATGCTCATTTGTGGAGAGCTTCATCATGGGCATA 600  
Db 92 TGCTCGCTCCAGGAGAGAGCCTTTGAAATGCTCATTCGTTGGGAGCCTTCATCATGGGCATA 151  
Qy 601 GATGAGAGCGCTCGCCCTCAGTGTCTCCCATGTACCTCAGTGTGAGATCTCACCAAGGAG 660  
Db 152 GATGAGAGCGCTCGCCCTCAGTGTCTCCCATGTACCTCAGTGTGAGATCTCACCAAGGAG 211  
Qy 661 ATCCGTGGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCACATGGG 720  
Db 212 ATCCGTGGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAATTTGGCGTGTTCACATGGG 271  
Qy 721 CAGCTTCTGGGCTGCTGGGAGAGAGTACCTGGGAGAGTACCTGGGAGAGTACCTGGGAGAG 780  
Db 272 CAGCTTCTGGGCTGCTGGGAGAGAGTACCTGGGAGAGTACCTGGGAGAGTACCTGGGAGAG 331  
Qy 781 GTGATTTGTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 332 GTGATTTGTGCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 391  
Qy 841 CGCTACCTGCTCTTGGAGAGAGCAACAGAGGAGAGAGTCTGTAAGAGCTTCCAAACGTTTC 900  
Db 392 CGCTACCTGCTCTTGGAGAGAGCAACAGAGGAGAGAGTCTGTAAGAGCTTCCAAACGTTTC 451  
Qy 901 TTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGAGGTCTCTGGCTGAGAGCCAGTGCAG 960  
Db 452 TTGGGTAAAGCAGACATTTCCCAAGAGGTAGAGAGGTCTCTGGCTGAGAGCCAGTGCAG 511  
Qy 961 AGGAGCATCCGCTGGTGTCCGCTGCTGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAG 1020  
Db 512 AGGAGCATCCGCTGGTGTCCGCTGCTGAGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAG 571  
Qy 1021 GTGCTACCGTGAATGTCAATGGCTGCTACAGAGCTGTGGCTCAATGCAATTTGG 1080  
Db 572 GTGCTACCGTGAATGTCAATGGCTGCTACAGAGCTGTGGCTCAATGCAATTTGG 631  
Qy 1081 TTCTATACCAAGAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 1140  
Db 632 TTCTATACCAAGAGCATCTTTGGAAAGCTGGATCCCTCCGGCAAGATCCCATACGTC 691  
Qy 1141 ACCTTGAGTACAGGGGGCATCGAGACTTTGGGTGCGCTCTTCTCTGGTTTGGTCAATTGAG 1200  
Db 692 ACCTTGAGTACAGGGGGCATCGAGACTTTGGGTGCGCTCTTCTCTGGTTTGGTCAATTGAG 751  
Qy 1201 CACCTGGAGCGGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 1260  
Db 752 CACCTGGAGCGGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGG 811



Db 337 CTCTGGTCTGTGACTGTGCATATTCGCCATCGTGGACTTGTGGGACGTTAAATGTG 396  
QY 457 AAGATGATTCGAAAGGTTCTTGGAGGAGACACACTTTGCTGCCAATATGTTTGA 516  
Db 397 AAGATGATTCGAAAGGTTCTTGGAGGAGACACACTTTGCTGCCAATATGTTTGA 456  
QY 517 ATTTCTGCTGATTCGATGAGGCTGCTCGCTCCAGGAGGAGCCTTTGAAATGCTCAIT 576  
Db 457 ATTTCTGCTGATTCGATGAGGCTGCTCGCTCCAGGAGGAGCCTTTGAAATGCTCAIT 516  
QY 577 GTGGGAGCCTTCATCATGGGCATAGATGGAGGCTGCTCCCTCAGTGCTGCCCATGTAC 636  
Db 517 GTGGGAGCCTTCATCATGGGCATAGATGGAGGCTGCTCCCTCAGTGCTGCCCATGTAC 576  
QY 637 CTCAGTGAGATCTCACCCAGGAGATCGTGGCTCTCTGGGAGGAGTGCATGCTTT 696  
Db 577 CTCAGTGAGATCTCACCCAGGAGATCGTGGCTCTCTGGGAGGAGTGCATGCTTT 636  
QY 697 ATCTGCATTGGCGTTTCACTGGGAGCTTCTGGGCTGCCGAGCTGCTGGGAGGAG 756  
Db 637 ATCTGCATTGGCGTTTCACTGGGAGCTTCTGGGCTGCCGAGCTGCTGGGAGGAG 696  
QY 757 AGTACCTGGCCATACCTGTTTGGAGTGAATGTGCTGCCCTGCTGCTGCTGCTGAGC 816  
Db 697 AGTACCTGGCCATACCTGTTTGGAGTGAATGTGCTGCCCTGCTGCTGCTGAGC 756  
QY 817 CTTCCTCTTCTCCCGGACAGCCACGCTACCTGCTTTGGAGAGCAACAGGAGCA 876  
Db 757 CTTCCTCTTCTCCCGGACAGCCACGCTACCTGCTTTGGAGAGCAACAGGAGCA 816  
QY 877 GCTGTGAAGCCTTCCAAACGTTCTTGGGTAAAGCAGAGTTTCCCAAGAGTAGAGG 936  
Db 817 GCTGTGAAGCCTTCCAAACGTTCTTGGGTAAAGCAGAGTTTCCCAAGAGTAGAGG 876  
QY 937 GTCCTGCTGAGAGGACGCTGACAGGAGCATCGGCTGCTGCTGCTGCTGCTGCTG 996  
Db 877 GTCCTGCTGAGAGGACGCTGACAGGAGCATCGGCTGCTGCTGCTGCTGCTGCTG 936  
QY 997 AGAGTCCCTACGCTCGCTGCGAGTGTGACCTGATTTGTCACCATGGCTGCTACAG 1056  
Db 937 AGAGTCCCTACGCTCGCTGCGAGTGTGACCTGATTTGTCACCATGGCTGCTACAG 996  
QY 1057 CTCTGTGCTCTCAATGCAATTTGTTCTATACCAAGAGCATTTTGAAGAGCTGGATC 1116  
Db 997 CTCTGTGCTCTCAATGCAATTTGTTCTATACCAAGAGCATTTTGAAGAGCTGGATC 1056  
QY 1117 CCTCGGCAAGATCCCATACGTCACCTTGAGTACAGGGGACATCGAGCTTTGGCTGCC 1176  
Db 1057 CCTCGGCAAGATCCCATACGTCACCTTGAGTACAGGGGACATCGAGCTTTGGCTGCC 1116  
QY 1177 GTCCTCTGCTGTTGGTCAATGAGCAGCTGGAGAGACCCCTCCTCATTTGGTGGCTTT 1236  
Db 1117 GTCCTCTGCTGTTGGTCAATGAGCAGCTGGAGAGACCCCTCCTCATTTGGTGGCTTT 1176  
QY 1237 GGGCTCATGGGCTCTTCTTTGGAGCCCTCACCATCGCTGACCTTCGAGGACACGCC 1296  
Db 1177 GGGCTCATGGGCTCTTCTTTGGAGCCCTCACCATCGCTGACCTTCGAGGACACGCC 1236  
QY 1297 CCCTGGGTCCCTACCTGAGTATCGTGGGCAATTCGCGCATCATCGCTCTTTCTGAGT 1356  
Db 1237 CCCTGGGTCCCTACCTGAGTATCGTGGGCAATTCGCGCATCATCGCTCTTTCTGAGT 1296  
QY 1357 GGGCCAGGTGGCATCC 1373  
Db 1297 GGGCCAGGTGTTTCCC 1313

RESULT 11  
US-10-188-186-85  
; Sequence 85, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:

; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO: 85  
; LENGTH: 1189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)..(1189)  
US-10-188-186-85

Query Match 48.4%; Score 1133.8; DB 12; Length 1189;  
Best Local Similarity 99.4%; Pred. No. 1.2e-301;  
Matches 1138; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 481 AGGAAGACACATTTGCTGCGCAATAATGGGTTTGCAATTTCTGCTGCAATGCTGATGCC 540  
Db 32 AGGAAGACACATTTGCTGCGCAATAATGGGTTTGCAATTTCTGCTGCAATGCTGATGCC 91  
QY 541 TGCTCGCTCCAGGAGGAGCCTTTGAAATGCTCATTTGGGACGCTTCATCATGAGGCATA 600  
Db 92 TGCTCGCTCCAGGAGGAGCCTTTGAAATGCTCATTTGGGACGCTTCATCATGAGGCATA 151  
QY 601 GATGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG 660  
Db 152 GATGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG 211  
QY 661 ATCCGTGGCTCTCTGGGACGCTGACTGCCATCTTTATCTGCAATGGGCTGTTCACTGGG 720  
Db 212 ATCCGTGGCTCTCTGGGACGCTGACTGCCATCTTTATCTGCAATGGGCTGTTCACTGGG 271  
QY 721 CAGCTTCTGGGCTGCGCCGAGCTGCTGGGAAAGGAGAGTACCTGGCCATACCTGTTTGA 780  
Db 272 CAGCTTCTGGGCTGCGCCGAGCTGCTGGGAAAGGAGAGTACCTGGCCATACCTGTTTGA 331  
QY 781 GTGATTTGGTCCCTGCGCTTGTGAGCTGCTGAGCTTCCCTTTCTCCGAGCAGCCCA 840  
Db 332 GTGATTTGGTCCCTGCGCTTGTGAGCTGCTGAGCTTCCCTTTCTCCGAGCAGCCCA 391  
QY 841 CGCTACCTGCTCTTGGAGAGCAGCAGGAGGAGCTGTGAAAGCCTTCCAAAGCTTC 900  
Db 392 CGCTACCTGCTCTTGGAGAGCAGCAGGAGGAGCTGTGAAAGCCTTCCAAAGCTTC 451  
QY 901 TTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGGCTCCTGGCTGAGAGCCACGTCAG 960  
Db 452 TTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGGCTCCTGGCTGAGAGCCACGTCAG 511

[illegible]

	Query Match	48.7%;	Score 1141;	DB 14;	Length 2216;
	Best Local Similarity	99.1%;	Pred. No. 1.7e-303;		
	Matches 1147;	Conservative	0;	Mismatches 10;	Indels 0; Gaps 0
Qy	217	AGAAAGACATGGCTCTCGCTCGCTCTCGTGGCCCTCCCTCGGGGGCGCCTTCGGGTCTCC	276		
Db	157	AAATTGGACTGGTCTCTCGCTCGCTCTCGTGGCTTCTCTCGGGGGCGCCTTCGGGTCTCC	216		
Qy	277	TTCTCTTCAACGGGTACAAACCTGTGGTGGTGAATGCCGCCACCCCGGTACATCAAGGCCCTTT	336		
Db	217	TTCTCTTCAACGGGTACAAACCTGTGGTGGTGAATGCCGCCACCCCGGTACATCAAGGCCCTTT	276		
Qy	337	TACAATGAGTCATGGGAAAAGAGGCATGGAGCTGCTCAATAGACCAGACACTCTGACTCTG	396		
Db	277	TACAATGAGTCATGGGAAAAGAGGCATGGAGCTGCTCAATAGACCAGACACTCTGACTCTG	336		
Qy	207	CTATCCATCCTGCACTGCTGCTCTCCATTTCTGCGCATCTGGTGGACTTTGGGGACGCTTAATCTGT	456		

	Query Match	48.7%;	Score 1141;	DB 14;	Length 2216;	
	Best Local Similarity	99.1%;	Pred. No. 1.7e-303;			
	Matches 1147;	Conservative	0;	Mismatches 10;	Indels	Gaps 0
Qy	217	AGAAAGACATGGCTCGTCCGTCCTCGTGCCTCCTCGGGGCGCGCTTCGGGTGCCCTCC	276			
Dd	157	AAATTGACTGSGTCTCTGCTCCTCGTGCCCTTCCTCGGGGCGCGCTTCGGGTGCCCTCC	216			
Qy	277	TTCCTCTACGGGTACAACCTGTCCGTGGTGAATGCCCGCCCACTCAATCAAAGCCCTTT	336			
Dd	217	TTCCTCTACGGGTACAACCTGTCCGTGGTGAATGCCCGCCCACTCAATCAAAGCCCTTT	276			
Qy	337	TACAATGAGTCATGGGAAAAGAAGCATGGAGCTGCAATAGACCAGACACTCTGACTCTG	396			
Dd	277	TACAATGAGTCATGGGAAAAGAAGCATGGAGCTGCAATAGACCAGACACTCTGACTCTG	336			
	207	CTATCCATCCTGCACTGCTGCTCCATAATTCCGCAATCGGTGGACTTTGGGGACGTTAATCTGTG	456			

1081 CACCTTGAGTACAGGGGCATCGAGACTTTGGCTGCCGTCTTCTC----- 1125  
1200 GCACCTGGGAGCGGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGCCCTCTTCTTTGG 1259  
1126 ----- 1125  
1260 GACCCCTCACCATCAGCTGACCCCTGCAGGACCAACGCCCTTGGGTCCCTTACCTGAGTAT 1319  
1126 ----- 1125  
1320 COTGGGCATTTCTGGCCATCATCGCCTCTTTCTGCAGTGGGCCAGGTGGCATCCCGTTTCAT 1379  
1126 -----TGGCATCCCGTTTCAT 1140  
1380 CTTGACTGGTGGTCTTCCAGCAATCTCAGGGCCGGCTGCCCTTCATCATTCGAGGCAC 1439  
1141 CTTGACTGGTGGTCTTCCAGCAATCTCAGGGCCGGCTGCCCTTCATCATTCGAGGCAC 1200  
1440 CGTCAACTGGCTCTCCAACTTTGCTTGGGTCTCTCTTCCCATTCATTCAAGAAAGTCT 1499  
1201 CGTCAACTGGCTCTCCAACTTTGCTTGGGTCTCTCTTCCCATTCATTCAAGAAAGTCT 1260  
1500 GGACACCTACTGTTTCTTCTAGTCTTTGTACAAATTTGTATCACAGTGTCTATCTACTCTGA 1559  
1261 GGACACCTACTGTTTCTTCTAGTCTTTGTACAAATTTGTATCACAGTGTCTATCTACTCTGA 1320  
1560 TTTTGTGTCTGCTGAGACCAAAACAGAACCTATGCAGAAATCAGCCAGGCAATTTTCCAA 1619  
1321 TTTTGTGTCTGCTGAGACCAAAACAGAACCTATGCAGAAATCAGCCAGGCAATTTTCCAA 1380  
1620 AAGGAACAAAGCATACCCACAGAACAGAAATCGACTCAGCTGTCTACTGTATGCT 1574  
1381 AAGGAACAAAGCATACCCACAGAACAGAAATCGACTCAGCTGTCTACTGTATGCT 1435

RESULT 9

US-10-188-186-89  
; Sequence 89, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO: 89  
; LENGTH: 1267  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:

1 GTCACTGAGACCCATGSCAAGGAAACAAAATAGAAATTTCAAGGAACCTGGGCCCTAGTTCC 60  
120 CCTCACAGATGACACCAAGCCACGCGGGCCCTCCAGGGCCAGGGAGGCGCACTGCTGGAGTG 179  
61 CCTCACAGATGACACCAAGCCACGCGGGCCCTCCAGGGCCAGGGAGGCGCACTGCTGGAGTG 120  
180 TGACCACTGAGAGTGGGTGCGAGGTGAAGAGGAGAAAGCACTGCTGCTGCTGCT 239  
121 TGTCCACTGAGGAGTGGGTGCGAGGTGAAGAGGAGAAAGCACTGCTGCTGCTGCTGCT 180  
240 CTTGCTGGCTCCTCTGCGGGCGCTTCCGCTCCTCTCTCTCTACGGCTACCAACCTGTC 299  
181 CTTGCTGGCTCCTCTGCGGGCGCTTCCGCTCCTCTCTCTCTACGGCTACCAACCTGTC 240  
300 GGTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTCAATGAGTCAATGGGAAAGAAAG 359  
241 GGTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTCAATGAGTCAATGGGAAAGAAAG 300  
360 GCATGAGCGTCCCAATAGACCCAGACACTCTGACTCTGCTCTGCTCTGCTCTGCTCTGCT 419  
301 GCATGAGCGTCCCAATAGACCCAGACACTCTGACTCTGCTCTGCTCTGCTCTGCTCTGCT 360  
420 ATTGCCCATCGGTGGACCTTGTGGGAGCTTAATTTGAAGATGATTTGGAAAGTCTTCTGG 479  
361 ATTGCCCATCGGTGGACCTTGTGGGAGCTTAATTTGAAGATGATTTGGAAAGTCTTCTGG 420  
480 GAGGAAGCACTTTGCTGGCAATATGCTCAATTTGGGAGCGCTTCAATCATGCGGCAT 539  
421 GAGGAAGCACTTTGCTGGCAATATGCTCAATTTGGGAGCGCTTCAATCATGCGGCAT 480  
540 CTGCTGCTCAGGACGAGCGCTTTGAAATGCTCAATTTGGGAGCGCTTCAATCATGCGGCAT 599  
481 CTGCTGCTCAGGACGAGCGCTTTGAGATGCTCATCTGCGGAGCGCTTCAATCATGCGGCAT 540  
600 AGATGAGGCGCTCGCCCTCAGTGTGCTCCCATGTACTCAGTGAATCTTCAACCAAGGA 659  
541 AGATGAGGCGCTCGCCCTCAGTGTGCTCCCATGTACTCAGTGAATCTTCAACCAAGGA 600  
660 GATCGGTGCTCTCTGCGGAGGAGTGCATGCTTATCTGATGAGTGGGTGCTGCTGCTGCT 719  
601 GATCGGTGCTCTCTGCGGAGGAGTGCATGCTTATCTGATGAGTGGGTGCTGCTGCTGCTG 660  
720 GCAGCTTCTGGGCTGCGGAGTGCCTGGGAAAGGAGTACCTGGCCATACCTGTTGG 779  
661 GCAGCTTCTGGGCTGCGGAGTGCCTGGGAAAGGAGTACCTGGCCATACCTGTTGG 720  
780 AGTGAATGCTGCTGCTGCGGTGCTCAGCTGCTCAGGCTTCCCTTTCTCCCGGACAGCC 839  
721 AGTGAATGCTGCTGCTGCGGTGCTCAGCTGCTCAGGCTTCCCTTTCTCCTGGACAGCC 780  
840 ACGCTACTGCTCTTGGAGAGCAACAGAGGCAAGAGCTGTGAAGCGCTTCCAAAGCTT 899  
781 ACGCTACTGCTCTTGGAGAGCAACAGAGGCAAGAGCTGTGAAGCGCTTCCAAAGCTT 840  
900 CTTGGGTAAGCAGACGTTTCCCAAGAGGTAGAGGAGTCTGCTGAGAGCCACGTCGA 959  
841 CTTGGGTAAGCAGACGTTTCCCAAGAGGTAGAGGAGTCTGCTGAGAGCCGCGTCGA 900  
960 GAGGAGCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1019  
901 GAGGAGCATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960  
1020 GGTGCTACCCGTGATTTGACCATGGGCTGCTACAGCTCTGTGGCCTCAATGCAATTTG 1079  
961 GGTGCTACCCGTGATTTGACCATGGGCTGCTACAGCTCTGTGGCCTCAATGCAATTTG 1020  
1080 GTTCTATACCAACAGCATTTTGGAAAGCTGGGATCCCTCCGCGCAAGATCCCATACGT 1139  
1021 GTTCTATACCAACAGCATTTTGGAAAGCTGGGATCCCTCTGGCAAGATCCCATACGT 1080  
1140 CACCTTGATGACAGGGGCGCATCGAGCTTTGGCTGCGCTCTTCTCTGTTGGTCAATTGA 1199

Db 241 TCACAGGTGGGCAAGCTTAGCCGCCAGAAAGTCAAGTCACTTGCTCAGACTCCTCAGCT 300  
QY 337 -----TACAATGAGTCATGGGAAAGAGCATGGACGTCCTCAATAGACCCAGACACTCTGA 391  
Db 301 GAGGGGACTGGCCCTGGAGGTAAAGCTGATATCATCTTGGCTCAAAGCCCAAGCTCTAT 360  
QY 392 CTCTCTCTGG-----TCTGTGACTGTGTCCATTAATTCGCC 426  
Db 361 CTCGTGGCTGGTGGCACTAGAGGAGACAAACGAGATTGGCAGAGACTGGTCTCTGGCTGCT 420  
QY 427 ATCGGTGGACATTGTGGGACGTTAATTGTG-----AAGATGAT 464  
Db 421 COTCGTGCCCTCCCTCGCGGGCGCTTCGGCTCCTCCTCTACCGCTACAACTGTGC 480  
QY 465 TGGAAAGTTCCTTGGGAGGAAGCACACTTTGCTGGCCCAATAATAGGTTTGCATTTCTGC 524  
Db 481 GGTGTGTAATGCCCCCAACCCCGCACACTTTGCTGGCCCAATAATAGGTTTGCATTTCTGC 540  
QY 525 TSCATTGTGATGGCTCTCGCTCCAGGAGGAGCCTTTGAAATGCTCATTTGTGGAGG 584  
Db 541 TGCAATTGCTGATGGCTCTCGCTCCAGGAGGAGCCTTTGAAATGCTCATCTGTGGAGG 600  
QY 585 CTTTCATCATGGGCATAGATGAGGCGTGGCCCTCAGTGTGCTCCCATGTACCTCAGTGA 644  
Db 601 CTTTCATCATGGGCATAGATGAGGCGTGGCCCTCAGTGTGCTCCCATGTACCTCAGTGA 660  
QY 645 GATCTCACCAAGAGATCCGTGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAT 704  
Db 661 GATCTCACCAAGAGATCCGTGCTCTCTGGGGCAGGTGACTGCCATCTTTATCTGCAT 720  
QY 705 TGGCGTGTCACTGGGAGCTTTGGGGCTGCGCCGAGCTGCTGGGAAAGAGAGTACCTG 764  
Db 721 TGGCGTGTCACTGGGAGCTTTGGGGCTGCGCCGAGCTGCTGGGAAAGAGAGTACCTG 780  
QY 765 GCCATACCTGTTTGGAGTGATTGGTCCCTGCTGCGTTGTCCAGCTGCTGAGCCTTCCCTT 824  
Db 781 GCCATACCTGTTTGGAGTGATTGGTCCCTGCTGCGTTGTCCAGCTGCTGAGCCTTCCCTT 840  
QY 825 TCTCCCGAGACCCCACTGCTCTCTGGGAGACACAACAGGAGCAAGACTGTGAA 884  
Db 841 TCTCTGGACACCCCACTGCTCTCTGGGAGACACAACAGGAGCAAGACTGTGAA 900  
QY 885 AGCCTTCCAAAGCTTTTGGGTAAAGCAGACGTTTCCAGAGGTAGAGAGTCTCTGCG 944  
Db 901 AGCCTTCCAAAGCTTTTGGGTAAAGCAGACGTTTCCAGAGGTAGAGAGTCTCTGCG 960  
QY 945 TGAGAGCCACTGCAGAGGACATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004  
Db 961 TGAGAGCCGCTGCAGAGGACATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020  
QY 1005 CTACGTCCGCTGGCAGGTGGTCAACCGTGAATGTCAACCATGGCCTGCTTACAGCTCTGTGG 1064  
Db 1021 CTACGTCCGCTGGCAGGTGGTCAACCGTGAATGTCAACCATGGCCTGCTTACAGCTCTGTGG 1080  
QY 1065 CCTCAATGCAATTTGGTTCTATACCAAGCAGATCTTTGGAAAGCTGGGATCCCTCGGC 1124  
Db 1081 CCTCAATGCAATTTGGTTCTATACCAAGCAGATCTTTGGAAAGCTGGGATCCCTCGGC 1140  
QY 1125 AAAGATCCCATACGTCACCTTGAATACAGGGGGCATCGAGACTTTGGCTCCGCTCTCTC 1184  
Db 1141 AAAGATCCCATACGTCACCTTGAATACAGGGGGCATCGAGACTTTGGCTCCGCTCTCTC 1200  
QY 1185 TGGTTTGGTCAATGAGCACTTGGAGCGAGACCCCTCTCATTTGGTGGCTTTGGGCTCAT 1244  
Db 1201 TGGTTTGGTCAATGAGCACTTGGAGCGAGACCCCTCTCATTTGGTGGCTTTGGGCTCAT 1260  
QY 1245 GGGCTCTCTTTGGGACCTTCAACATCAGCTGACCTGACCTGAGGACCAACGCCCTCGGT 1304  
Db 1261 GGGCTCTCTTTGGGACCTTCAACATCAGCTGACCTGACCTGAGGACCAACGCCCTCGGT 1320  
QY 1305 CCCCTTACCTGATTCGTGGGCACTTCTGGCCCATCATCGCTCTCTTCTGAGTGGGCGAG 1364  
Db 1321 CCCCTTACCTGATTCGTGGGCACTTCTGGCCCATCATCGCTCTCTTCTGAGTGGGCGAG 1380

QY 1365 TGGCATCCCGTTCACTTGAAGTCTGGTGAAGTCTTCTCCAGCAATCTCAGCGCGGCTGCTT 1424  
Db 1381 TGGCATCCCGTTCACTTGAAGTCTGGTGAAGTCTTCTCCAGCAATCTCAGCGCGGCTGCTT 1440  
QY 1425 CATCATTCGAGGACCGCTCAACTGGCTCTCCAACTTTGCTGTTGGGCTCTCTCTCCATT 1484  
Db 1441 CATCATTCGAGGACCGCTCAACTGGCTCTCCAACTTTGCTGTTGGGCTCTCTCTCCATT 1500  
QY 1485 CATTCAGAAAAGTCTGGACACCTACTGTTTCTCTAGTCTTTTGTCTACAAATTTGTATCAGG 1544  
Db 1501 CATTCAGAAAAGTCTGGACACCTACTGTTTCTCTAGTCTTTTGTCTACAAATTTGTATCAGG 1560  
QY 1545 TGCTATCTACCTGTTATTTTGTGCTGCTGAGACCAAAACAGAACCTATTCAGAAATCAG 1604  
Db 1561 TGCTATCTACCTGTTATTTTGTGCTGCTGAGACCAAAACAGAACCTATTCAGAAATCAG 1620  
QY 1605 CCAGGCAATTTTCCAAAAGGAAACAAAGCATACCCACAGAGAGAAATCGACTCAGCTGT 1664  
Db 1621 CCAGGCAATTTTCCAAAAGGAAACAAAGCATACCCACAGAGAGAAATCGACTCAGCTGT 1680  
QY 1665 CACTGATGCT 1674  
Db 1681 CACTGATGCT 1690

RESULT 8  
US-10-188-186-79  
; Sequence 79, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Angerson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305673  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 79  
; LENGTH: 1502  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (14)..(1454)  
US-10-188-186-79

Query Match 52.5%; Score 1229; DB 12; Length 1502;  
Best Local Similarity 88.2%; Pred. No. 0;  
Matches 1425; Conservative 0; Mismatches 10; Indels 180; Gaps 1;  
QY 60 GTCACTGAGACCCCATGCGAGGAAACAAAATAGGAATTCAGAGAACTGGGCTAGTTCC 119

Qy 482 GGAAGCAGCATTGCTGGCCCAATAATGGGTTTGCATTTCTGCTGATTTGCTGATGGCT 541  
Db 358 GGAAGCAGCATTGCTGGCCCAATAATGGGTTTGCATTTCTGCTGATTTGCTGATGGCT 417  
Qy 542 GCTCGCTCCAGGAGAGAGCCTTTGAATGCTATTTGAGAGCCTTTCAATCATGGCCATAG 601  
Db 418 GCTCGCTCCAGGAGAGAGCCTTTGAATGCTATTTGAGAGCCTTTCAATCATGGCCATAG 477  
Qy 602 ATGAGAGCGTGGCCCTCAGTGTGCTCCCATGTACCTCAGTGTGAGATCTCACCAAGAGAGA 661  
Db 478 ATGAGAGCGTGGCCCTCAGTGTGCTCCCATGTACCTTAGTGTGAGATCTCACCAAGAGAGA 537  
Qy 662 TCCGTGCTCTCTGGGGCAGGTGACTGGCCATCTTTATCTGCAATTTGCGGTGTTCACTGGCC 721  
Db 538 TCCGTGCTCTCTGGGGCAGGTGACTGGCCATCTTTATCTGCAATTTGCGGTGTTCACTGGCC 597  
Qy 722 AGCTTCTGGGCTGGCCGAGCTGCTGGAAAGAGAGATACCTGGCCATACCTGTTTGGAG 781  
Db 598 AGCTTCTGGGCTGGCCGAGCTGCTGGAAAGAGAGATACCTGGCCATACCTGTTTGGAG 657  
Qy 782 TGATTTGGTCCCTGGCGTTGTCCAGTCTGTGAGCCTTCCCTTTCTCCGGACAGCCAC 841  
Db 658 TGATTTGGTCCCTGGCGTTGTCCAGTCTGTGAGCCTTCCCTTTCTCCGGACAGCCAC 717  
Qy 842 GCTACCTGCTCTGGAGAGACAAGAGAGAGAGCTGTGAAAGCTTCCAAAGCTTCT 901  
Db 718 GCTACCTGCTCTGGAGAGACAAGAGAGAGAGCTGTGAAAGCTTCCAAAGCTTCT 777  
Qy 902 TGGGTAAGCAGACGTTTCCCAAGAGAGAGAGAGCTGTGAGAGCTCCCTACGTCGCTGGCAGG 961  
Db 778 TGGGTAAGCAGACGTTTCCCAAGAGAGAGAGAGCTGTGAGAGCTCCCTACGTCGCTGGCAGG 837  
Qy 962 GGAGCATTCGCTGGTTCGCTGCTGAGCTGTGAGAGCTCCCTACGTCGCTGGCAGG 1021  
Db 838 GGAGCATTCGCTGGTTCGCTGCTGAGCTGTGAGAGCTCCCTACGTCGCTGGCAGG 897  
Qy 1022 TGGTACCGGTGATGTCACCATGGCTGTACAGCTCTGTGGCTCTCAATGCAATTTGGT 1081  
Db 898 TGGTACCGGTGATGTCACCATGGCTGTACAGCTCTGTGGCTCTCAATGCAATTTGGT 957  
Qy 1082 TCTATACCAAGCAGCTCTTTGAAAAGCTGGGATCCCTCCGCAAGAGATCCCATACGTCA 1141  
Db 958 TCTATACCAAGCAGCTCTTTGAAAAGCTGGGATCCCTCCGCAAGAGATCCCATACGTCA 1017  
Qy 1142 CTTTGAGTACAGGGGAGCATGAGACTTTGGCTGCGCTCTTCTGCTGTTGGTCAATGAGC 1201  
Db 1018 CTTTGAGTACAGGGGAGCATGAGACTTTGGCTGCGCTCTTCTGCTGTTGGTCAATGAGC 1077  
Qy 1202 ACCTGGAGCGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTCTTTGGGA 1261  
Db 1078 ACCTGGAGCGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTCTTTGGGA 1137  
Qy 1262 CCCTCACCATACAGCTGACCTGAGAGACAGCCCTGCTGGCTCCCTACCTGAGTATCG 1321  
Db 1138 CCCTCACCATACAGCTGACCTGAGAGACAGCCCTGCTGGCTCCCTACCTGAGTATCG 1197  
Qy 1322 TGGGCAATCTGCGCATCATGCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTCACT 1381  
Db 1198 TGGGCAATCTGCGCATCATGCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTCACT 1257  
Qy 1382 TGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTGCTTCATATGAGAGACCG 1441  
Db 1258 TGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTGCTTCATATGAGAGACCG 1317  
Qy 1442 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCATTCATTCAGAAAGTCTGG 1501  
Db 1318 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCATTCATTCAGAAAGTCTGG 1377  
Qy 1502 ACACCTACTGTTTCTTAGTCTTTGCTACAAATTTGATACAGAGTCTATCTACCTGTATT 1561  
Db 1378 ACACCTACTGTTTCTTAGTCTTTGCTACAAATTTGATACAGAGTCTATCTACCTGTATT 1437

Qy 1562 TTGTGCTGCTGAGACCAAAAACAGAACCTATGCGAATAATCAGCCAGGCAATTTTCCAAA 1621  
Db 1438 TTGTGCTGCTGAGACCAAAAACAGAACCTATGCGAATAATCAGCCAGGCAATTTTCCAAA 1497  
Qy 1622 GGAACAAAACATACCCACCAGAGAGAAATCGACTGTCTACTGATGCT 1674  
Db 1498 GGAACAAAACATACCCACCAGAGAGAAATCGACTGTCTACTGATGCT 1550

RESULT 7  
US-10-188-186-77  
; Sequence 77, Application US/10188186  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502  
; PRIOR FILING DATE: 2001-07-11  
; PRIOR APPLICATION NUMBER: 60/305262  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: 60/373881  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 60/305873  
; PRIOR FILING DATE: 2001-07-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 368  
; SOFTWARE: Custom  
; SEQ ID NO 77  
; LENGTH: 1757  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (14)..(1709)  
; US-10-188-186-77

Query Match 52.6%; Score 1232.8; DB 12; Length 1757;  
Best Local Similarity 85.7%; Pred. No. 0;  
Matches 1448; Conservative 0; Mismatches 167; Indels 75; Gaps 4;  
Qy 60 GTCACTGAGACCCATGGCAGGAGAAACAAATAGGAATTCAGAGAACTGGGCTGATTTCC 119  
Db 1 GTCACTGAGACCCATGGCAGGAGAAACAAATAGGAATTCAGAGAACTGGGCTGATTTCC 60  
Qy 120 CCTCAGATGACACAGCCAGCCGCGGCTCCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGG 179  
Db 61 CCTCAGATGACACAGCCAGCCGCGGCTCCAGGCGCAGGAGGAGGAGGAGGAGGAGGAGG 120  
Qy 180 TGACCACCTGAGGAGTGGGCTGCCAGGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 234  
Db 121 TGACCACCTGAGGAGTGGGCTGCCAGGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180  
Qy 235 TGGCTCCTGCTGGCTCCCTCGCGGGGCGCTTGGGCTCCTCCTTCTCTACGGCTACAC 294  
Db 181 CTCCTCTGCGAGGCGCTCCTCAACAAACATATGTGCCAGTGTCTGCTATTAGATCCCAT 240  
Qy 295 CTGTGCTGTGATGTCGCCCGCCAGCCCGGTACATCAAGGCTTT----- 336

Db 418 GCTCGCTCCAGGAGGAGCCCTTTGAAATGCTCATCGTGGGACGCTTCATCATGGGCATAG 477  
Qy 602 ATGAGGCGTCGCGCTCAGTGTCTCCCATGTACCTCAGTGAGATCTCACCAGGAGA 661  
Db 478 ATGAGGCGTCGCGCTCAGTGTCTCCCATGTACCTCAGTGAGATCTCACCAGGAGA 537  
Qy 662 TCGGTGGCTCTCTGGGCGAGGTGACTGCCATCTTTATCTGCAATGGCGTGTCACTGGGC 721  
Db 538 TCGGTGGCTCTCTGGGCGAGGTGACTGCCATCTTTATCTGCAATGGCGTGTCACTGGGC 597  
Qy 722 AGCTTCTGGGCTGCGCGAGTCTCTGGGAAAGAGAGTACCTGGCCATACCTGTTGGAG 781  
Db 598 AGCTTCTGGGCTGCGCGAGTCTCTGGGAAAGAGAGTACCTGGCCATACCTGTTGGAG 657  
Qy 782 TGATTGTGCTGCGCGTGTCTGAGCTGTGAGCCTTCCCTTTCTCCGGACAGCCAC 841  
Db 658 TGATTGTGCTGCGCGTGTCTGAGCTGTGAGCCTTCCCTTTCTCCGGACAGCCAC 717  
Qy 842 GCTACCTGCTCTTGGAGAGACACACAGGCGAAGAGCTGTGAAAGCTTCCAAAGCTTCT 901  
Db 718 GCTACCTGCTCTTGGAGAGACACACAGGCGAAGAGCTGTGAAAGCTTCCAAAGCTTCT 777  
Qy 902 TGGGTAAGCAGACGTTTCCCAAGAGGTAGAGAGTCTGCGCTGAGAGCCACGTCGAGA 961  
Db 778 TGGGTAAGCAGACGTTTCCCAAGAGGTAGAGAGTCTGCGCTGAGAGCCGCGTCGAGA 837  
Qy 962 GAGCATCCGCTGTGTCTGCTGTGAGCTGTGAGAGTCCCTACGTCGCTGGCAGG 1021  
Db 838 GAGCATCCGCTGTGTCTGCTGTGAGCTGTGAGAGTCCCTACGTCGCTGGCAGG 897  
Qy 1022 TGGTACCGTGTATGACCATGGCTGTACAGCTCTGTGGCTCAATGCAATTTGGT 1081  
Db 898 TGGTACCGTGTATGACCATGGCTGTACAGCTCTGTGGCTCAATGCAATTTGGT 957  
Qy 1082 TCTATACCAACAGCATCTTTGAAAGCTGGATCCCTCCGGAAGATCCCATACGTCA 1141  
Db 958 TCTATACCAACAGCATCTTTGAAAGCTGGATCCCTCCGGAAGATCCCATACGTCA 1017  
Qy 1142 CCTTGAGTACAGGGGATCGAGACTTTGGCTCGCTCTTCTCTGCTTGTGTCATTGAGC 1201  
Db 1018 CCTTGAGTACAGGGGATCGAGACTTTGGCTCGCTCTTCTCTGCTTGTGTCATTGAGC 1077  
Qy 1202 ACTGGGACGAGACACCCCTCTCATGTTGGTCTTTGGGCTCATGGGCTCTTCTTTGGGA 1261  
Db 1078 ACTGGGACGAGACACCCCTCTCATGTTGGTCTTTGGGCTCATGGGCTCTTCTTTGGGA 1137  
Qy 1262 CCTCACCATCAGCTGACCTTCGAGGACACGCGCCCTGGGTCCCTACCTGAGTATCG 1321  
Db 1138 CCTCACCATCAGCTGACCTTCGAGGACACGCGCCCTGGGTCCCTACCTGAGTATCG 1197  
Qy 1322 TGGGCATTCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGTGGCATCCGTTCACT 1381  
Db 1198 TGGGCATTCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGTGGCATCCGTTCACT 1257  
Qy 1382 TGACTGTTGAGTCTTCCAGCAATCTCAGCGGCGGCTCGCTTTCATCATTTGAGGACCG 1441  
Db 1258 TGACTGTTGAGTCTTCCAGCAATCTCAGCGGCGGCTCGCTTTCATCATTTGAGGACCG 1317  
Qy 1442 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCTTCCCAATTCATTGAGGAGGCTGCG 1501  
Db 1318 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCTTCCCAATTCATTGAGGAGGCTGCG 1377  
Qy 1502 ACACCTACTGTTTCTAGTCTTTGCTACAAATTTGATCAGAGTGTCTATCTACCTGATT 1561  
Db 1378 ACACCTACTGTTTCTAGTCTTTGCTACAAATTTGATCAGAGTGTCTATCTACCTGATT 1437  
Qy 1562 TTGTCGTCTGAGACCAACAAACACAACTATGAGAAATCAGCAGGCAATTTTCCAAA 1621  
Db 1438 TTGTCGTCTGAGACCAACAAACACAACTATGAGAAATCAGCAGGCAATTTTCCAAA 1497  
Qy 1622 GGAAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGTCTACCTGATGCT 1674  
Db 1498 GGAAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGTCTACCTGATGCT 1550

RESULT 6

US-10-117-722-574  
; Sequence 574, Application US/10117722  
; Publication No. US20030219744A1  
; GENERAL INFORMATION:  
; APPLICANT: Tang, Y. Tom  
; APPLICANT: Liu, Chenghua  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Zhang, Jie  
; APPLICANT: Dmanac, Radoje T.  
; TITLE OF INVENTION: No. US20030219744A1el Nucleic Acids and  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 784CIP2BCIP  
; CURRENT APPLICATION NUMBER: US/10/117,722  
; CURRENT FILING DATE: 2002-04-04  
; PRIOR APPLICATION NUMBER: 09/620,312  
; PRIOR FILING DATE: 2000-07-19  
; PRIOR APPLICATION NUMBER: 05/552,317  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/488,725  
; PRIOR FILING DATE: 2000-01-21  
; NUMBER OF SEQ ID NOS: 1104  
; SOFTWARE: pf\_fl\_genes Version 1.0  
; SEQ ID NO 574  
; LENGTH: 2842  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (345)..(1571)  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)..(2842)  
; OTHER INFORMATION: n = a,t,c or g  
US-10-117-722-574

Query Match 56.2%; Score 1317.8; DB 15; Length 2842;  
Best Local Similarity 89.9%; Pred.No. 0;  
Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 2;  
Qy 2 CGACCCACGCGTCGCGCTTTGGCAGAGTCTGGGGTCCCTGGACTGAGGCCATCAGCTGGGT 61  
Db 40 CGACCCACGCGTGC-GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGGCCATCAGCTGGGT 98  
Qy 62 CACTGAGACCCATGGCAAGGAAACAAATAGGAATTCAGGAACTGGGCTAGTTCCTCC 121  
Db 99 CACTGAGACCCATGGCAAGGAAACAAATAGGAATTCAGGAACTGGGCTAGTTCCTCC 158  
Qy 122 TCACAGATGACACAGGCCAGCCGGGCTCCAGGGCCAGGGAGGCACTGCTGGAGTGTG 181  
Db 159 TCACAGATGACACAGGCCAGCCGGGCTCCAGGGCCAGGGAGGCACTGCTGGAGTGTG 218  
Qy 182 ACCACCTGAGGAGTGGGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 241  
Db 219 ACCACCTGAGGAGTGGGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 278  
Qy 242 TCGTGGCTCCCTCGGGGGCGCTTCGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGG 301  
Db 279 TCGTGGCTCCCTCGGGGGCGCTTCGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCGG 338  
Qy 302 TGTGAATGCCCCACCCCGTACATCAAGGCTTTTACAATAGATCATGGGAAAGAGGC 361  
Db 339 TGTGAATGCCCCACCCCGTACATCAAGGCTTTTACAATAGATCATGGGAAAGAGGC 357  
Qy 362 ATGGAGCTCCATAGAGCCAGACACTCTGACTCTGCTCTGCTCTGCTCTGCTCTGCTCATAT 421  
Db 358 ----- 357  
Qy 422 TCGCCATCGGTGAGCTTTGGGGAGCGTTAATTGTAAGATGATTGGAAAGGTTCTTTGGGA 481  
Db 358 ----- 357



Db 622 AGTGATTGTTGGTCCCTGCGGTTGTCAGCTGCTGAGCCTTCCCTTTCTCCCGGACAGCCC 681  
Qy 840 AGCTACCTGCTCTTGGAGAAGCACAACGAGGCAAGAGCTGTGAAGGCTTCCAAACGTT 899  
Db 682 AGCTACCTGCTCTTGGAGAAGCACAACGAGGCAAGAGCTGTGAAGGCTTCCAAACGTT 741  
Qy 900 CTTGGTAAAGCAGAGCTTTCCAGAGCTAGAGAGTCTGAGGCTGAGGCCAGCTGCA 959  
Db 742 CTTGGTAAAGCAGAGCTTTCCAGAGCTAGAGAGTCTGAGGCTGAGGCCAGCTGCA 801  
Qy 960 GAGGAGCATCCGCTGGTGTCCGCTGCTGAGAGTCTGAGAGTCTCCCTACGTCGCTGGCA 1019  
Db 802 GAGGAGCATCCGCTGGTGTCCGCTGCTGAGAGTCTGAGAGTCTCCCTACGTCGCTGGCA 861  
Qy 1020 GTTGCTACCGGTATGTACCATGCTGCTACAGCTCTGACAGCTCTGTGCTCAATGCAATTTG 1079  
Db 862 GTTGCTACCGGTATGTACCATGCTGCTACAGCTCTGTGCTCAATGCAATTTG 921  
Qy 1080 GTTCTATACCAACAGCATCTTTGAAAAGCTGGATCCCTCGGCAAAAGATCCCATAGT 1139  
Db 922 GTTCTATACCAACAGCATCTTTGAAAAGCTGGATCCCTCGGCAAAAGATCCCATAGT 981  
Qy 1140 CACCTTGCTAGCAGGGGATCGAGACTTTGGCTGCGCTCTCTCTGTTTGGTCATTGA 1199  
Db 982 CACCTTGCTAGCAGGGGATCGAGACTTTGGCTGCGCTCTCTCTGTTTGGTCATTGA 1041  
Qy 1200 GCACTTGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTGG 1259  
Db 1042 GCACTTGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTGG 1101  
Qy 1260 GACCTTACCATCAGCTGACCTGACGAGCAACGCGCCCTGGGTCCCTACCTGAGTAT 1319  
Db 1102 GACCTTACCATCAGCTGACCTGACGAGCAACGCGCCCTGGGTCCCTACCTGAGTAT 1161  
Qy 1320 CGTGGGCTAGTGGGCTATCGCTCTTTTGGAGTGGGCTGGGCTGGGCTGGGCTGGGCT 1379  
Db 1162 CGTGGGCTAGTGGGCTATCGCTCTTTTGGAGTGGGCTGGGCTGGGCTGGGCTGGGCT 1221  
Qy 1380 CTTGACTGTGAGTCTTCCAGCAATCTCAGGCGCGGCTGCTTCATCATTTGAGGCAAC 1439  
Db 1222 CTTGACTGTGAGTCTTCCAGCAATCTCAGGCGCGGCTGCTTCATCATTTGAGGCAAC 1281  
Qy 1440 CGTCAACTGGCTCTCCAACTTTGGCTGTTGGGCTCTCTTCCCATTCATTCAGAAAAGTCT 1499  
Db 1282 CGTCAACTGGCTCTCCAACTTTGGCTGTTGGGCTCTCTTCCCATTCATTCAGAAAAGTCT 1341  
Qy 1500 GGACACCTACTGTTCTCTAGTCTTTGCTACAAATTTGTATCAGAGTGTATCTACCTGTA 1559  
Db 1342 GGACACCTACTGTTCTCTAGTCTTTGCTACAAATTTGTATCAGAGTGTATCTACCTGTA 1401  
Qy 1560 TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATTCAGCAGGCAATTTCCAA 1619  
Db 1402 TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATTCAGCAGGCAATTTCCAA 1461  
Qy 1620 AAGGAACAAAGCATACCCACCAAGAGAGAAATCGACTGCTGCTGATGCT 1674  
Db 1462 AAGGAACAAAGCATACCCACCAAGAGAGAAATCGACTGCTGCTGATGCT 1516

RESULT 5

US-10-037-270-574

Sequence 574, Application US/10037270

Publication No. US20030104529A1

GENERAL INFORMATION:

APPLICANT: Tang, Y. Tom

APPLICANT: Liu, Chenghua

APPLICANT: Asundi, Vinod

APPLICANT: Zhang, Jie

APPLICANT: Ren, Feiyan

APPLICANT: Chen, Rui-hong

APPLICANT: Zhao, Qing A.

APPLICANT: Wehrman, Tom

APPLICANT: Xue, Aldong J.

APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yuning  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: Tillinghast, John  
APPLICANT: Dmanac, Radoje T.  
TITLE OF INVENTION: No. US20030104529A1el Nucleic Acids and  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/10/037, 270  
CURRENT FILING DATE: 2002-01-04  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1104  
SOFTWARE: Pt\_FL\_genes Version 1.0  
SEQ ID NO 574  
LENGTH: 2842  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: CDS  
LOCATION: (345)..(1571)  
NAME/KEY: misc feature  
LOCATION: (1)..(2842)  
OTHER INFORMATION: n = a,t,c or g  
US-10-037-270-574

Query Match 56.2%; Score 1317.8; DB 14; Length 2842;

Best Local Similarity 89.9%; Pred. No. 0;

Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 2;

Qy 2 CGACCCACGGCTCCGGCTTTGGCAGAGTCTGGGGTCCCTGGAGTACGACATGAGTGGGT 61  
Db 40 CGACCCACGGCTGC-GCCTTTGGCAGAGTCTGGGGTCCCTGGAGTACGACATGAGTGGGT 98  
Qy 62 CACTGAGACCCATGGCAAGAAACAAATAGGAATTCAGAGGACTGGGCTAGTTCCTCC 121  
Db 99 CACTGAGACCCATGGCAAGAAACAAATAGGAATTCAGAGGACTGGGCTAGTTCCTCC 158  
Qy 122 TCACAGATGACACAGCAGCGCGGCTCCAGGGCCAGGGAGGAGGACCTGCTGGAGTGTG 181  
Db 159 TCACAGATGACACAGCAGCGCGGCTCCAGGGCCAGGGAGGAGGACCTGCTGGAGTGTG 218  
Qy 182 ACCACCTGAGGAGTGGGGTCCAGAGTGGAGGAGAGAGGAGGAGGAGGAGGAGGAGGAGG 241  
Db 219 ACCACCTGAGGAGTGGGGTCCAGAGTGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 278  
Qy 242 TCGTGGCTCCCTCGGGGGCGCTTCGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 301  
Db 279 TCGTGGCTCCCTCGGGGGCGCTTCGGGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 338  
Qy 302 TGGTGAATGCCCCCAGCCCGGTACATCAAGGCTTTTACATGAGTATGGAAGAGGAGGAGG 361  
Db 339 TGGTGAATGCCCCCAGCCCGG 357  
Qy 362 ATGAGCTCCATAGACCCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
Db 358 357  
Qy 422 TCGCCATCGGTGGACTTGTGGGAGCGTTAATTGTGAAGATGATTGGAAGGTTCTTGGGA 481  
Db 358 357  
Qy 482 GGAAGCAGACTTTGCTGGCCAAATAATGGGTTTCCAAATTTCTGCTGATGCTGATGCTGCT 541  
Db 358 GGAAGCAGACTTTGCTGGCCAAATAATGGGTTTCCAAATTTCTGCTGATGCTGATGCTGCT 417  
Qy 542 GCTGCTCCAGGAGGAGGCTTTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 601

RESULT 4  
US-10-188-186-75  
; Sequence 75, Application US/10188-86  
; Publication No. US20040029789A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 21402-397C  
; CURRENT APPLICATION NUMBER: US/10/188,186  
; CURRENT FILING DATE: 2002-07-02  
; PRIOR APPLICATION NUMBER: 60/303046  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/360814  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/303828  
; PRIOR FILING DATE: 2001-09-07  
; PRIOR APPLICATION NUMBER: 60/323380  
; PRIOR FILING DATE: 2001-09-19  
; PRIOR APPLICATION NUMBER: 60/361133  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: 60/304016  
; PRIOR FILING DATE: 2001-07-09  
; PRIOR APPLICATION NUMBER: 60/304502

262	ATTCGCAUCGGTGGACTTTGTGGGACATTAATTGAGAGATGATGGAAAGGTTCTTGG	321
480	GAGGAAGCACACTTTGCTGCCCAATAATGGGTTTGCAATTTCTGCTGCATTTGCTGATGC	539
322	GAGGAAGCACACTTTGCTGCCCAATAATGGGTTTGCAATTTCTGCTGCATTTGCTGATGC	381
540	CTGCTCGCTCCAGGACAGGACCTTTGAAATGCTCATTTGTGGGACGCTTCATCATGGGCAT	599
382	CTGCTCGCTCCAGGACAGGACCTTTGAAATGCTCATCGTGGACGCTTCATCATGGGCAT	441
600	AGATGAGGCGTCGCCCTCAGTGTGCTCCCCATGTACCTCAGTGAGATCTCACCCCAAGA	659
442	AGATGAGGCGTCGCCCTCAGTGTGCTCCCCATGTACCTCAGTGAGATCTCACCCCAAGA	501
660	GATCCGTGCTCTCTGGGCGAGTGACTGCCATCTTTATCTGCATTTGGCGTGTCACTGG	719
502	GATCCGTGCTCTCTGGGCGAGTGACTGCCATCTTTATCTGCATTTGGCGTGTCACTGG	561
720	GCAGCTTCTGGGCGTCCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACTGTTTGG	779
562	GCAGCTTCTGGGCGTCCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACTGTTTGG	621
780	AGTGAATTGTGGTCCCTGCGGTGTCCACGCTGCTGAGCCTTCCCTTTCTCCCGACAGCCC	839



841	Db		AGAGCAACAGGCGAAGAGCTGTGAAAGCCTTCCAAACGTTCTTGGTTAAAGCAGACG	900
917	Qy		TTTCCCAAGAGGTAGAGAGGTCTCGCTGAGAGCAGGTGCAGAGGAGCATCGCTGG	976
901	Db		TTTCCCAAGAGGTAGAGAGGTCTCGCTGAGAGCGGTGCAGAGGAGCATCGCTCG	960
977	Qy		TGTCGCTGCTGGAGCTGCTGAGAGTCCCTACGTCCGCTGGCAGGTGTCACCGTGATTG	1036
961	Db		TGTCGCTGCTGGAGCTGCTGAGAGTCCCTACGTCCGCTGGCAGGTGTCACCGTGATTG	1020
1037	Qy		TCACCATGGCTGCTACAGCTCTGTGGCTCAATGCAATTGGTTCTATACCAACAGCA	1096
1021	Db		TCACCATGGCTGCTACAGCTCTGTGGCTCAATGCAATTGGTTCTATACCAACAGCA	1080
1097	Qy		TCCTTGGAAAAGCTGGGATCCCTCCGCGAAAGATCCCATCGTCACCTTCAGTACAGGGG	1156
1081	Db		TCCTTGGAAAAGCTGGGATCCCTCTGCGAAAGATCCCATCGTCACCTTCAGTACAGGGG	1140
1157	Qy		GCATCGAGACTTGGCTGCCGCTTCTCTGTTTGGTCAATTGACGACCTCGGACGAGAC	1216
1141	Db		GCATCGAGACTTGGCTGCCGCTTCTCTGTTTGGTCAATTGAGCACCTGGGACGAGAC	1200
1217	Qy		CCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGGACCCCTCACCATCAGC	1276
1201	Db		CCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGGGACCCCTCACCATCAGC	1260
1277	Qy		TGACCCCTGCAGGACCAAGCCCTCGGCTCCCTACCTACGATATCGTGGGCAATTCTGGCCA	1336
1261	Db		TGACCCCTGCAGGACCAAGCCCTCGGCTCCCTACCTACGATATCGTGGGCAATTCTGGCCA	1320
1337	Qy		TCATCGGCTCTTTCTGCGAGTGGGCCAGGTGGCATCCCGTTTCATCTTCAGCTGGTGAATTCT	1396
1321	Db		TCATCGGCTCTTTCTGCGAGTGGGCCAGGTGGCATCCCGTTTCATCTTCAGCTGGTGAATTCT	1380
1397	Qy		TCCAGCAATCTCAGCGCGGCTGCTTCATCATTGACAGGCACCGTCAACTGGCTCTCCA	1456
1381	Db		TCCAGCAATCTCAGCGCGGCTGCTTCATCATTGACAGGCACCGTCAACTGGCTCTCCA	1440
1457	Qy		ACTTTGCTGTGGGCTCCCTCTCCCATTCATTCAGAAAAGTCTGGACACCTACTGTTTCC	1516
1441	Db		ACTTTGCTGTGGGCTCCCTCTCCCATTCATTCAGAAAAGTCTGGACACCTACTGTTTCC	1500
1517	Qy		TAGCTTTTGCTACAAATTGTATCACAGGTGCTATCTACTGTATTTTGTGCTGCCGTGAGA	1576
1501	Db		TAGCTTTTGCTACAAATTGTATCACAGGTGCTATCTACTGTATTTTGTGCTGCCGTGAGA	1560
1577	Qy		CCAAAAACAGAACTATGCAGAAATCAGCCAGGCATTTTCCAAAGGAAACAAAGCATACC	1636
1561	Db		CCAAAAACAGAACTATGCAGAAATCAGCCAGGCATTTTCCAAAGGAAACAAAGCATACC	1620
1637	Qy		CACCAGAAAGAAAATCGACTCAGCTGCTCATGTGCTC	1675
1621	Db		CACCAGAAAGAAAATCGACTCAGCTGCTCATGTGCTC	1659

```

RESULT 3
US-10-276-774-930
; Sequence 930. Application US/10276774
; Publication No. US20040053245A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc.
; APPLICANT: Tang, Y, Tom et al
; TITLE OF INVENTION: No. US20040053245A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 21272-030
; CURRENT APPLICATION NUMBER: US/10/276,774
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: 09/560,875
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: 09/496,914
; PRIOR FILING DATE: 2000-02-03
; NUMBER OF SEQ ID NOS: 2700
; SOFTWARE: Custom

```

QY	1981	AGAGAGAGAGAGAGATGCTATTCTAAAAAGCTTGAGTCTAGGCTGTGCACGGTGGCTC	2040
Db	1981	AGAGAGAGAGAGAGAGATGCTATTCTAAAAAGCTTGAGTCTAGGCTGTGCACGGTGGCTC	2040
QY	2041	ACGCCCTTAATCCCGACACTTTGGAGGCGGAGGTGGGTGGATCGTAGGTCAAGGAGATT	2100
Db	2041	ACGCCCTTAATCCCGACACTTTGGAGGCGGAGGTGGGTGGATCGTAGGTCAAGGAGATT	2100
QY	2101	GAGACCATTCTGGCTAAACATGGTGAACAACCTCCCTCTCTACTTAAAAAATACAAAAATTAGCT	2160
Db	2101	GAGACCATTCTGGCTAAACATGGTGAACAACCTCCCTCTCTACTTAAAAAATACAAAAATTAGCT	2160
QY	2161	GAGCATGGTGGCGGCGCCCTGTAGTCCCAGCTACTTTGGGAGGCTGAGGCGAGGAGANTGGC	2220
Db	2161	GAGCATGGTGGCGGCGCCCTGTAGTCCCAGCTACTTTGGGAGGCTGAGGCGAGGAGANTGGC	2220
QY	2221	GTGAACCCAGAGGCGGAGCTTGCAGTGAAGCCGAGATCACACCACACTCCAGCCTGG	2280
Db	2221	GTGAACCCAGAGGCGGAGCTTGCAGTGAAGCCGAGATCACACCACACTCCAGCCTGG	2280
QY	2281	GTGACAGAGCCAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGGC	2340
Db	2281	GTGACAGAGCCAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGGC	2340
QY	2341	CGC 2343	
Db	2341	CGC 2343	

RESULT 2  
US-10-297-022-52  
; Sequence 52, Application US/10297022  
; Publication No. US20030216310A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.  
; APPLICANT: THORNTON, Michael  
; APPLICANT: WALIA, Narinder K.  
; APPLICANT: YUE, Henry  
; APPLICANT: NGUYEN, Danniell B.  
; APPLICANT: LAL, Preeti  
; APPLICANT: GANDHI, Ameena R.  
; APPLICANT: TRIBOULEY, Catherine M.  
; APPLICANT: YAO, Monique G.  
; APPLICANT: RAMKUMAR, Jayalaxmi  
; APPLICANT: AU-YOUNG, Janice  
; APPLICANT: LU, Yan  
; APPLICANT: TANG, Y. Tom  
; APPLICANT: AZIMZAI, Yalda  
; APPLICANT: BRUNS, Christopher M.  
; APPLICANT: GRIFFIN, Jennifer A.  
; APPLICANT: YANG, Junning  
; APPLICANT: BAUGHN, Mariah R.  
; APPLICANT: SANJANWALA, Madhu S.  
; APPLICANT: RAUMANN, Brigitte E.  
; APPLICANT: LEE, Ernestine A.  
; APPLICANT: HAPALIA, April  
; APPLICANT: GREENE, Barrie D.  
; APPLICANT: KHAN, Farrah A.  
; APPLICANT: KEARNEY, Liam  
; APPLICANT: ELLIOTT, Vicky S.  
; APPLICANT: SEITHAMER, Jeffrey J.  
; APPLICANT: POLICKY, Jennifer L.  
; APPLICANT: BOROWSKI, Mark L.  
; APPLICANT: BURFORD, Neil  
; APPLICANT: DING, Li  
; APPLICANT: LU, Dyung Aina M.  
; APPLICANT: HILLMAN, Jennifer L.  
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS  
; FILE REFERENCE: PI-0109 PCT  
; CURRENT APPLICATION NUMBER: US/10/297,022  
; PRIOR FILING DATE: 2002-11-25  
; PRIOR APPLICATION NUMBER: 60/208,424  
; 60/209,001; 60/210,588; 60/212,335; 60/213,747;

## RESULT 2

US-10-297-022-52

Sequence 52, Application US/10297022

Publication No. US20030216310A1

GENERAL INFORMATION:

APPLICANT: INCYTE GENOMICS, INC.

APPLICANT: THORNTON, Michael

APPLICANT: WALIA, Narinder K.

APPLICANT: YUE, Henry

APPLICANT: NGUYEN, Dannel B.

APPLICANT: LAL, Preci

APPLICANT: GANDHI, Ameena R.

APPLICANT: TRIBOULEY, Catherine M.

APPLICANT: YAO, Monique G.

APPLICANT: RAMKUMAR, Jayalaxmi

APPLICANT: AU-YOUNG, Janice

APPLICANT: LU, Yan

APPLICANT: TANG, Y. Tom

APPLICANT: AZIMZAI, Yalda

APPLICANT: BRUNS, Christopher M.

APPLICANT: GRIFFIN, Jennifer A.

APPLICANT: YANG, Junming

APPLICANT: BAUGHN, Mariah R.

APPLICANT: SANJANWALA, Madhu S.

APPLICANT: RAUMANN, Brigitte E.

APPLICANT: LEE, Ernestine A.

APPLICANT: HAFALIA, April

APPLICANT: GREENE, Bartle D.

APPLICANT: KHAN, Farrah A.

APPLICANT: KEARNEY, Liam

APPLICANT: ELLIOTT, Vicky S.

APPLICANT: SEILHAMER, Jeffrey J.

APPLICANT: POLICKY, Jennifer L.

APPLICANT: BOROWSKY, Mark L.

APPLICANT: BURFORD, Neil

APPLICANT: DING, Li

APPLICANT: LU, Dyung Aina M.

APPLICANT: HILLMAN, Jennifer L.

TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS

FILE REFERENCE: PI-0109 PCT

CURRENT APPLICATION NUMBER: US/10/297,022

CURRENT FILING DATE: 2002-11-25

PRIOR APPLICATION NUMBER: 60/208,424; 60/209,001; 60/210,588; 60/212,335; 60/213,747

LENGTH: 2343 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 73...1761  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-981-947A-1

Query Match 100.0%; Score 2343; DB 9; Length 2343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCAGACCACGCGTCCGGCCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG 60  
Db 1 TCAGACCACGCGTCCGGCCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG 60

Qy 61 TCACGTGAGCCCATGGCAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTTCCTC 120  
Db 61 TCACGTGAGCCCATGGCAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTTCCTC 120

Qy 121 CTCACAGATACACACAGCCAGCCGCGGGCCCTCAGGCCCAAGGAGGCACTGCTGAGTGT 180  
Db 121 CTCACAGATACACACAGCCAGCCGCGGGCCCTCAGGCCCAAGGAGGCACTGCTGAGTGT 180

Qy 181 GACCACCTGAGGAGTGGGGTCCAGGTGGAGGAGAAAGAGTGTCTGCTCGCTC 240  
Db 181 GACCACCTGAGGAGTGGGGTCCAGGTGGAGGAGAAAGAGTGTCTGCTCGCTC 240

Qy 241 CTCGTGCTCCTCCGCGGGCGCTTCGGCTCCTCCTCTACGGCTACAACCTGTCG 300  
Db 241 CTCGTGCTCCTCCGCGGGCGCTTCGGCTCCTCCTCTACGGCTACAACCTGTCG 300

Qy 301 GTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTACATGAGTCAATGGAAAGAGG 360  
Db 301 GTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTACATGAGTCAATGGAAAGAGG 360

Qy 361 CATGGACGTCATAGACACAGACACTCTGACTCTGCTCTGCTGCTGACTGTGTCATA 420  
Db 361 CATGGACGTCATAGACACAGACACTCTGACTCTGCTCTGCTGCTGACTGTGTCATA 420

Qy 421 TTGCCCATCGGTGGACTGTGGGAGCGTTAAATGTGAAGATGATTTGGAAGGTTCTTGG 480  
Db 421 TTGCCCATCGGTGGACTGTGGGAGCGTTAAATGTGAAGATGATTTGGAAGGTTCTTGG 480

Qy 481 AGGAAGACACTTGTGGGCAATATGGGTTTGGCAATTTCTGCTGCTGCTGCTGATGGCC 540  
Db 481 AGGAAGACACTTGTGGGCAATATGGGTTTGGCAATTTCTGCTGCTGCTGCTGATGGCC 540

Qy 541 TGCTCGCTCCAGACAGAGCGCTTGAATGCTCATTTGGGAGCGCTTCATCATGGGCATA 600  
Db 541 TGCTCGCTCCAGACAGAGCGCTTGAATGCTCATTTGGGAGCGCTTCATCATGGGCATA 600

Qy 601 GATGGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAAAGGAG 660  
Db 601 GATGGAGGCGTCCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAAAGGAG 660

Qy 661 ATCGTGGCTCTCTGGGCGAGTGAATGCCATCTTATCTGCAATTTGGCGGTGCTACTGG 720  
Db 661 ATCGTGGCTCTCTGGGCGAGTGAATGCCATCTTATCTGCAATTTGGCGGTGCTACTGG 720

Qy 721 CAGCTTCTGGGCTGCCGAGCTGTGGAAAGGAGAGTACCTGGCCATACCTGTTTGA 780  
Db 721 CAGCTTCTGGGCTGCCGAGCTGTGGAAAGGAGAGTACCTGGCCATACCTGTTTGA 780

Qy 781 GTGATTTGGTCTCCGCTGTGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840  
Db 781 GTGATTTGGTCTCCGCTGTGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 840

Qy 841 CGTACTGCTCTTGGAGAGCACAACAGGAGGAGAGTGTGGAAGCTTCCAAAGCTTC 900  
Db 841 CGTACTGCTCTTGGAGAGCACAACAGGAGGAGAGTGTGGAAGCTTCCAAAGCTTC 900

Db 841 CGCTACTGCTCTTTGGAGAAAGCACACAGGAGCAAGAGCTGTGAAAGCTTCCAAAGCTTC 900  
Qy 901 TTGGGTAAAGCAGACAGCTTTCCCAAGAGGTAGAGGAGTCCCTGGCTGAGAGCCACCTGCAG 960  
Db 901 TTGGGTAAAGCAGACAGCTTTCCCAAGAGGTAGAGGAGTCCCTGGCTGAGAGCCACCTGCAG 960

Qy 961 AGGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
Db 961 AGGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020

Qy 1021 GTGGTCACCGTGTATTTGTACACCATGGCCCTGCTACACAGCTCTGTGGCTCAATGCAATTTGG 1080  
Db 1021 GTGGTCACCGTGTATTTGTACACCATGGCCCTGCTACACAGCTCTGTGGCTCAATGCAATTTGG 1080

Qy 1081 TTCTATACAAACAGCATCTTTTGGAAAGCTGGGATCCCTCCGGCAAAAGATCCCATACCTC 1140  
Db 1081 TTCTATACAAACAGCATCTTTTGGAAAGCTGGGATCCCTCCGGCAAAAGATCCCATACCTC 1140

Qy 1141 ACCTTCACTACAGGGGCAATCGAGACTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
Db 1141 ACCTTCACTACAGGGGCAATCGAGACTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200

Qy 1201 CACCTGGGACGAGAGCCCTCTCTCATTTGGTGGCTTTGGCTCATGGGCTCTCTCTTTGGG 1260  
Db 1201 CACCTGGGACGAGAGCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTCTTTGGG 1260

Qy 1261 ACCCTCACCATCAACGCTGACCTCGAGGACACGCGCCCTCGGCTCCCTACTCCTGAGTATC 1320  
Db 1261 ACCCTCACCATCAACGCTGACCTCGAGGACACGCGCCCTCGGCTCCCTACTCCTGAGTATC 1320

Qy 1321 GTGGGCAATTTCTGGCCATCATCGCTCTTTTCTGAGTGGGCGAGGCTGCTGCTGCTGCTGCTG 1380  
Db 1321 GTGGGCAATTTCTGGCCATCATCGCTCTTTTCTGAGTGGGCGAGGCTGCTGCTGCTGCTGCTG 1380

Qy 1381 TTGACTGTGAGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTCATTTGCGAGGACCC 1440  
Db 1381 TTGACTGTGAGTCTTCCAGCAATCTCAGCGGCGGCTGCTTCTCATTTGCGAGGACCC 1440

Qy 1441 GTCAACTGCTCTCCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500  
Db 1441 GTCAACTGCTCTCCAACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1500

Qy 1501 GACACCTACTGTTTCTAGTCTTCTGCTCAATTTGTATCAAGGCTGCTATCTGCTGAT 1560  
Db 1501 GACACCTACTGTTTCTAGTCTTCTGCTCAATTTGTATCAAGGCTGCTATCTGCTGAT 1560

Qy 1561 TTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATCAGCCAGGCAATTTTCCAA 1620  
Db 1561 TTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATCAGCCAGGCAATTTTCCAA 1620

Qy 1621 AGGAACAAAGCATACCCACAGAGAGAAATTCGACTCAGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
Db 1621 AGGAACAAAGCATACCCACAGAGAGAAATTCGACTCAGCTGCTGCTGCTGCTGCTGCTGCTG 1680

Qy 1681 TCTTCTCTCTTCACTCTCCGAAATACGCTGGAATTCAGGCTGCGGCAACCAACCGCC 1740  
Db 1681 TCTTCTCTCTTCACTCTCCGAAATACGCTGGAATTCAGGCTGCGGCAACCAACCGCC 1740

Qy 1741 ACCAAAAAGACACCCATTTGAAACGCTCATGTGGTATTTCTTCAACCTGGAATGACCT 1800  
Db 1741 ACCAAAAAGACACCCATTTGAAACGCTCATGTGGTATTTCTTCAACCTGGAATGACCT 1800

Qy 1801 TCCCTTATCTTCTCTCTGGAGAACACCAAGTCTATGATGTGAGACAGAGCTTGGATTT 1860  
Db 1801 TCCCTTATCTTCTCTCTGGAGAACACCAAGTCTATGATGTGAGACAGAGCTTGGATTT 1860

Qy 1861 TGGAGACATGGGTTTGAATTTCCAGTCAATCTCTTTTATTTTATTTTATTTTAAACAGT 1920  
Db 1861 TGGAGACATGGGTTTGAATTTCCAGTCAATCTCTTTTATTTTATTTTATTTTAAACAGT 1920

Qy 1921 ACTGACATGCTCCCATATGTTTATCCACCTGTTTATACAAATGGAGGAGAGAGAG 1980  
Db 1921 ACTGACATGCTCCCATATGTTTATCCACCTGTTTATACAAATGGAGGAGAGAGAG 1980

Sequence 39, Appl  
Sequence 57, Appl  
Sequence 696, Appl  
Sequence 861, Appl  
Sequence 307, Appl  
Sequence 23, Appl  
Sequence 47, Appl  
Sequence 75, Appl  
Sequence 473, Appl  
Sequence 135, Appl  
Sequence 135, Appl  
Sequence 2128, Appl  
Sequence 135, Appl  
Sequence 1294, Appl  
Sequence 135, Appl  
Sequence 1031, Appl  
Sequence 135, Appl  
Sequence 11, Appl  
Sequence 12, Appl  
Sequence 13, Appl  
Sequence 14, Appl  
Sequence 7, Appl  
Sequence 6, Appl  
Sequence 381, Appl  
Sequence 9, Appl  
Sequence 8, Appl  
Sequence 10, Appl  
Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-09-981-947A-1  
; Sequence 1, Application US/09981947A  
; Patent No. US20020164578A1  
; GENERAL INFORMATION:  
; APPLICANT: Tartaglia, Louis A.  
; Weng, Xun  
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
; GUTEX AND USES THEREOF  
; NUMBER OF SEQUENCES: 10  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows95  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/981,947A  
; FILING DATE: 18-Oct-2001  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 09/031,392  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weiklejohn, Ph.D., Anita L.  
; REGISTRATION NUMBER: 35,283  
; REFERENCE/DOCKET NUMBER: 07334/072002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617/542-5070  
; TELEFAX: 617/542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 8, 2004, 14:27:39 ; Search time 819 Seconds  
(without alignments)  
10732.203 Million cell updates/sec

Title: US-09-981-947B-1

Perfect score: 2343

Sequence: 1 tcgaccacgcgtccgcct.....aaaaaaaaagggcgccgc 2343

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2475585 seqs, 1875730760 residues

Total number of hits satisfying chosen parameters: 4951170

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*

- 1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*
- 2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*
- 3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*
- 4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*
- 5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*
- 6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*
- 7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*
- 8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*
- 9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*
- 10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*
- 11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*
- 12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*
- 13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*
- 14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*
- 15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*
- 16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*
- 17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*
- 18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2343	100.0	2343	9	US-09-981-947A-1
2	1651	70.5	2032	15	US-10-297-022-52
3	1644.6	70.2	1851	12	US-10-276-774-930
4	1399	59.7	1710	12	US-10-188-186-75
5	1317.8	56.2	2842	14	US-10-037-270-574
6	1317.8	56.2	2842	15	US-10-117-722-574
7	1232.8	52.6	1757	12	US-10-188-186-77
8	1229	52.5	1502	12	US-10-188-186-79
9	1187.6	50.7	1267	12	US-10-188-186-89
10	1141	48.7	2216	14	US-10-168-651-33
11	1133.8	48.4	1189	12	US-10-188-186-85
12	1130.6	48.3	1189	12	US-10-188-186-83
13	1130.6	48.3	1189	12	US-10-188-186-87
14	915	39.1	1087	12	US-10-188-186-81
15	430.8	18.4	472	10	US-09-918-995-30642

Search completed: April 8, 2004, 16:44:45  
Job time : 196 secs

OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20  
Query Match 11.7%; Score 273.2; DB 2; Length 246240;  
Best Local Similarity 86.3%; Pred. No. 1.7e-54;  
Matches 302; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
TYPE: nucleic acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1.246240  
OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-21  
Query Match 11.7%; Score 273.2; DB 2; Length 246240;  
Best Local Similarity 86.3%; Pred. No. 1.7e-54;  
Matches 302; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
QY 1983 AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTTAGGCTGTGACCGTGGCTCAC 2042  
DB 2949 AGAGATATAGATCTCTCTTTTATCTTAAAGCTTTTATAGGCGCGGCGGTGGCTCAC 3008  
QY 2043 GCCTGTATCCAGACACTTTGGAGCGCGAGTGGTGGATCGTGAAGTTCAGAGATTGA 2102  
DB 3009 GCCTGTATCCAGACACTTTGGAGCGCGAGTGGTGGATCGTGAAGTTCAGAGATTGA 3068  
QY 2103 GACCATCTGGCTAATCATGTGAACTCCCTCTCTACTTAAATAACAAAAATTAGCTGA 2162  
DB 3069 GACCATCTGGCTAATCATGTGAACTCCCTCTCTACTTAAATAACAAAAATTAGCTGA 3128  
QY 2163 GCATGTGGCGCGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGGAGAGATGGCGT 2222  
DB 3129 GCATGTGGCGCGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGGAGAGATGGCGT 3188  
QY 2223 GAACCCAGAGCGGAGCTTTGAGTGGAGCGAGATCACACACACACCTCCAGCTGGGT 2282  
DB 3189 GAACCCAGAGCGGAGCTTTGAGTGGAGCGAGATCACACACACACCTCCAGCTGGGT 3248  
QY 2283 GACAGAGCAGACTCCGCTCTCAAAAAAATAAAAAAATAAAAAA 2332  
DB 3249 GACAGAGCAGACTCCGCTCTCAAAAAAATAAAAAAATAAAAAA 3298  
RESULT 15  
US-08-724-394A-22  
Sequence 22, Application US/08724394A  
Patent No. 587237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 587237el  
TITLE OF INVENTION: Sequences and Antibodies Thereto  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136

OTHER INFORMATION: /note= "HLA-H.CONTIG"  
US-08-724-394A-20  
Query Match 11.7%; Score 273.2; DB 2; Length 246240;  
Best Local Similarity 86.3%; Pred. No. 1.7e-54;  
Matches 302; Conservative 0; Mismatches 48; Indels 0; Gaps 0;  
QY 1983 AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTTAGGCTGTGACCGTGGCTCAC 2042  
DB 2949 AGAGATATAGATCTCTCTTTTATCTTAAAGCTTTTATAGGCGCGGCGGTGGCTCAC 3008  
QY 2043 GCCTGTATCCAGACACTTTGGAGCGCGAGTGGTGGATCGTGAAGTTCAGAGATTGA 2102  
DB 3009 GCCTGTATCCAGACACTTTGGAGCGCGAGTGGTGGATCGTGAAGTTCAGAGATTGA 3068  
QY 2103 GACCATCTGGCTAATCATGTGAACTCCCTCTCTACTTAAATAACAAAAATTAGCTGA 2162  
DB 3069 GACCATCTGGCTAATCATGTGAACTCCCTCTCTACTTAAATAACAAAAATTAGCTGA 3128  
QY 2163 GCATGTGGCGCGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGGAGAGATGGCGT 2222  
DB 3129 GCATGTGGCGCGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGGAGAGATGGCGT 3188  
QY 2223 GAACCCAGAGCGGAGCTTTGAGTGGAGCGAGATCACACACACACCTCCAGCTGGGT 2282  
DB 3189 GAACCCAGAGCGGAGCTTTGAGTGGAGCGAGATCACACACACACCTCCAGCTGGGT 3248  
QY 2283 GACAGAGCAGACTCCGCTCTCAAAAAAATAAAAAAATAAAAAA 2332  
DB 3249 GACAGAGCAGACTCCGCTCTCAAAAAAATAAAAAAATAAAAAA 3298  
RESULT 14  
US-08-724-394A-21  
Sequence 21, Application US/08724394A  
Patent No. 587237  
GENERAL INFORMATION:  
APPLICANT: Feder, John N.  
APPLICANT: Krommal, Gregory S.  
APPLICANT: Lauer, Peter M.  
APPLICANT: Ruddy, David A.  
APPLICANT: Thomas, Winston  
APPLICANT: Tsuchihashi, Zenta  
APPLICANT: Wolff, Roger K.  
TITLE OF INVENTION: Megabase Transcript Map: No. 587237el  
TITLE OF INVENTION: Sequences and Antibodies Thereto  
NUMBER OF SEQUENCES: 31  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: 01-OCT-1996  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Fitts, Renee A.  
REGISTRATION NUMBER: 35,136  
REFERENCE/DOCKET NUMBER: 017957-000100  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300  
INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 246240 base pairs

RESULT 12  
 US-08-759-873-5/c  
 ; Sequence 5, Application US/08759873  
 ; Patent No. 5683885  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kieback, Dirk G.  
 ; TITLE OF INVENTION: METHODS FOR DIAGNOSING AN INCREASED RISK  
 ; TITLE OF INVENTION: OF BREAST OR OVARIAN CANCER  
 ; NUMBER OF SEQUENCES: 14  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
 ; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800  
 ; CITY: Washington, D.C.  
 ; STATE: D.C.  
 ; COUNTRY: U.S.A.  
 ; ZIP: 20037  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/759,873  
 ; FILING DATE: 12-APRIL-1996  
 ; CLASSIFICATION: 435  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: KIT, Gordon  
 ; REGISTRATION NUMBER: 30,764  
 ; REFERENCE/DOCKET NUMBER: A-6612  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: (202) 293-7060  
 ; TELEFAX: (202) 293-7860  
 ; INFORMATION FOR SEQ ID NO: 5:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 320 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA  
 ; HYPOTHETICAL: NO  
 US-08-759-873-5

Query Match 11.7%; Score 275.2; DB 1; Length 320;  
Best Local Similarity 91.2%; pred. No. 5e-56;

RESULT 13  
 US-08-724-394A-20  
 ; Sequence 20, Application US/08724394A  
 ; Patent No. 587237  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Feder, John N.  
 ; APPLICANT: Krommal, Gregory S.  
 ; APPLICANT: Lauer, Peter M.  
 ; APPLICANT: Ruddy, David A.  
 ; APPLICANT: Thomas, Winston  
 ; APPLICANT: Teuchihashi, Zenta  
 ; APPLICANT: Wolff, Roger K.  
 ; TITLE OF INVENTION: Megabase Transcript Map: No. 5872237el  
 ; TITLE OF INVENTION: Sequences and Antibodies Thereto  
 ; NUMBER OF SEQUENCES: 31  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: TOWNSEND and TOWNSEND and CREW LLP  
 ; STREET: Two Embarcadero Center, 8th Floor  
 ; CITY: San Francisco  
 ; STATE: CA  
 ; COUNTRY: USA  
 ; ZIP: 94111-3834  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/724.394A  
 ; FILING DATE: 01-OCT-1996  
 ; CLASSIFICATION: 536  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Fitts, Renee A.  
 ; REGISTRATION NUMBER: 35,136  
 ; REFERENCE/DOCKET NUMBER: 017957-000100  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 415-576-0200  
 ; TELEFAX: 415-576-0300  
 ; INFORMATION FOR SEQ ID NO: 20:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 246240 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: not relevant  
 ; TOPOLOGY: not relevant  
 ; MOLECULE TYPE: cDNA  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: 1..246240



Query Match 11.8%; Score 276.6; DB 4; Length 12619;  
Best Local Similarity 89.7%; Pred. No. 8.9e-56;  
Matches 297; Conservative 0; Mismatches 34; Indels 0; Gaps 0;  
Qy 2003 TTCTAAAGCTTGAAGTCTAGGCTGTGACGGTGGCTCAGCCTGTAAATCCAGCACTTT 2062  
Db 10625 TTTTAAAAATCTGCAGCATAGGCGCGGCGGTGCTCAGCCTGTAAATCCAGCACTTT 10566  
Qy 2063 GGGAGCCGAGGTGGTGGATCTGAGTCAAGAGATTGAGCATTGAGCCTGCTTAAATG 2122  
Db 10565 GGGAGCCGAGGTGGTGGATCTGAGTCAAGAGATTGAGCATTGAGCCTGCTTAAATG 10506  
Qy 2123 TGAACCTCCCTCTCTACTATAAATAACAAAAATAGCTGAGCATGGTGGCGGCGCTGT 2182  
Db 10505 TGAACCTCCCTCTCTACTATAAATAACAAAAATAGCTGAGCATGGTGGCGGCGCTGT 10446  
Qy 2183 AGTCCCAAGCTTACTGGAGGCTGAGCAGGAGATGGCGTGAACCCAGGAGGCGGCTT 2242  
Db 10445 AGTCCCAAGCTTACTGGAGGCTGAGCAGGAGATGGCGTGAACCCAGGAGGCGGCTT 10386  
Qy 2243 GCAGTGAAGCGAGATCACACCACTCCAGCTGGGTGACAGCAGCAGACTCCGCTCT 2302  
Db 10385 GCAGTGAAGCGAGATTCGCGCACTGCACTCGGCGGCGGCGGCGGCGGCGGCTCT 10326  
Qy 2303 CAAAAAATAAAAAAAAAAAAAAAAAAAAAA 2333  
Db 10325 CAAAAAATAAAAAAAAAAAAAAAAAAAAAA 10295

RESULT 11  
US-08-629-939-5/c  
; Sequence 5, Application US/08629939  
; Patent No. 5645995  
; GENERAL INFORMATION:  
; APPLICANT: Kieback, Dirk G.  
; TITLE OF INVENTION: METHODS FOR DIAGNOSING AN INCREASED  
; RISK OF BREAST OR OVARIAN CANCER  
; NUMBER OF SEQUENCES: 14  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SUGHRUE, MION, ZINN, MACPEAK & SEAS  
; STREET: 2100 Pennsylvania Avenue, N.W., Suite 800  
; CITY: Washington, D.C.  
; STATE: D.C.  
; COUNTRY: U.S.A.  
; ZIP: 20037  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/629,939  
; FILING DATE: 12-APRIL-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: KIT, Gordon  
; REGISTRATION NUMBER: 30,764  
; REFERENCE/DOCKET NUMBER: A-6612  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 293-7060  
; TELEFAX: (202) 293-7860  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 320 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA  
; HYPOTHETICAL: NO  
US-08-629-939-5

Query Match 11.7%; Score 275.2; DB 1; Length 320;  
Best Local Similarity 91.2%; Pred. No. 5e-56;

Qy 1018 CAGGTGGTCAACGATGTTACCATGGCTGCTACAGCTCTGGGCTCAATGCAATT 1077  
Db 990 CCATCCTCATCGCTGTGGTGTGAGCTGTCCAGCAGCTGTGCGCATCAACGCTGTC 1049  
Qy 1078 TGGTCTTACCAACAGCATCTTTGAAAAAGCTGGGATCCCTCCGCGAAAGATCCCATAC 1137  
Db 1050 TTCTATTACTCCAGCATCTTTGAGAAAGCGGGGGTGCAGCAGCTGTG-----TAT 1103  
Qy 1138 GTACACCTTGAGTACAGGGGAGCATCAGAGCTTTGGTGGCGCTTCTCTGTTTGGTCAAT 1197  
Db 1104 GCCACATTGGTCTCGGTATCGTCAACAGGCTTCACTGTGCTGCTGTTTGGTGTG 1163  
Qy 1198 GAGCAGCTGGGAGGAGACCTCTCTCATGTGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1257  
Db 1164 GAGCAGCAGGCGCGGAGACCTTGACCTCATAGGCTCGTGGCATGGCGGCTTGTGCC 1223  
Qy 1258 GGGAGCCTCAGCATCAGCTGAGCCCTGAGGACAGGCGGCTGGTGGTGGTGGTGGTGGT 1317  
Db 1224 ATACTCATGACCATCGCTAGCAGCTGCTGGAGCAGTACCTGGATGCTCTATCTGAGC 1283  
Qy 1318 ATCGTGGGATCTGCGCATATCGCTCTTTCTGAGTGGCGGAGTGGATCCCGTTC 1377  
Db 1284 ATCGTGGGATCTTTGGCTTTTGGCTTTTGAAGTGGTCTTGGCGCCATGCCATGG 1343  
Qy 1378 ATCTTGACTGGTGGTCTTCCAGCAATCTCAGCGCGGCTGCTTCCATTCATTTGAGGC 1437  
Db 1344 TTCTCGTGGTGAATCTTTCAGCAGGCTCCAGCTCCAGCTGCCATGCGGTTGAGGC 1403  
Qy 1438 ACCGTCACTGGCTTCCAACTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1497  
Db 1404 TTCTCACTGGAGCTCAAAATTCATTTGGGATGTGCTTCCAGTATGTGGAGCAATG 1463  
Qy 1498 CTGGACACTACTGTTTCTAGCTTTTGTGTAATTTGTATCAGGTTGCTATCTACCTG 1557  
Db 1464 TGTGGTCCCTAGCTTCTATCATCTTCACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523  
Qy 1558 TATTTTGTGCTGCTGAGCAACCAACAGCAACCTATGAGAAATC 1602  
Db 1524 TACTTCAAAAGTCTCCAGACTAAAGCCGAGCTTCGATGAGATC 1568

RESULT 10  
US-09-616-289-49/c  
; Sequence 49, Application US/09616289  
; Patent No. 6632923  
; GENERAL INFORMATION:  
; APPLICANT: Lees, Ann M.  
; APPLICANT: Lees, Robert S.  
; APPLICANT: Law, Simon W.  
; APPLICANT: Arjona, Anibal A.  
; TITLE OF INVENTION: NOVEL LOW DENSITY LIPOPROTEIN BINDING  
; PROTEINS AND THEIR USE IN DIAGNOSING AND TREATING  
; Atherosclerosis  
; FILE REFERENCE: 10797-004001  
; CURRENT APPLICATION NUMBER: US/09/616,289  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: US 09/517,849  
; PRIOR FILING DATE: 2000-03-02  
; PRIOR APPLICATION NUMBER: US 08/979,608  
; PRIOR FILING DATE: 1997-11-26  
; PRIOR APPLICATION NUMBER: US 60/031,930  
; PRIOR FILING DATE: 1996-11-27  
; PRIOR APPLICATION NUMBER: US 60/048,547  
; PRIOR FILING DATE: 1997-06-03  
; NUMBER OF SEQ ID NOS: 53  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 49  
; LENGTH: 12619  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-616-289-49



QY 538 GCCTGCTCGCTCCAGGACGAGCCTTTGAAATGCTCATTTGCGAGCGCTTCATCATGCGC 597  
Db 510 GGTCTTCTCGAAATGCGGCAAGTCTTTGAGATGCTCATCTCGGCGCTTCATCATCGGT 569  
QY 598 ATAGATGAGGCGTCCCTCAGTGTCTCCCATATGCTACCTCAGTGAATCTTCAACCCAAAG 657  
Db 570 GTGTACTTGGCGCTGACACAGGCTTCGTGCCCATGTATGTGGGTGAAGTGTCAACCCACA 629  
QY 658 GAGATCGGTGCTCTTGGGCGAGTGAAGTGCATCTTTATCTGATGTGGGTGTACT 717  
Db 630 GCCTTGTGGGCGCTTGGGCGAGTGCATCTTGTGAGTGTGGGCGATCTGCTGCGCATCTCATC 689  
QY 718 GGGCAGCTTCTGGGCGCTGCCGAGTGTGGGAAAGGAGTACCTGGCCATACCTGTTT 777  
Db 690 GCCCAGGTGTTCGGCTGAGTCTCATCTGAGTGGGCAAGAGGAGTGTGGCGCTGCTGTG 749  
QY 778 GAGATGATGTGCTTCCCTGCGTGTGCTGAGCTGTGAGCTTCCCTTCTTCCGAGACAGC 837  
Db 750 AGCATCATCTTATCCCGGCGCTGTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809  
QY 838 CCAAGCTTACCTGCTTGGGAGGACACAGGAGGCAAGAGTGTGAAGGCTTCCAAAG 897  
Db 810 CCGCGCTTCTGCTCATCAACCGCAACGAGGAGAACCGGCGCCAAAGTGTGCTTAAAGAG 869  
QY 898 TTCTTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGTCTTGGCTGAGAGCCACGTG 957  
Db 870 CTGCGGGACAGCTGACGTGACCTGACCTGAGGAGATGAAGGAGAGTGTGGCAG 929  
QY 958 CAGAGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1017  
Db 930 ATGATCGGAGAGAGAGTCAACATCTGAGAGTGTTCGCTGCTGCTGCTGCTGCTGCTGCTG 989  
QY 1018 CAGGTGCTACCGTGAATTCACCATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
Db 990 CCAATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1049  
QY 1078 TGGTCTTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGGCAAGATCCCATAC 1137  
Db 1050 TTCTATTACTCCACAGCATCTTCGAGAGGCGGGGTGCGAGCCTGTG-----TAT 1103  
QY 1138 GTCACTTGAATACAGGGGCGATGAGACTTTGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197  
Db 1104 GCCACATTTGGCTCCGATATCGTCAACAGCGCTTCACTGTGCTGCTGCTGCTGCTGCTGCTG 1163  
QY 1198 GAGCACTGGGAGAGACCTCTCTCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257  
Db 1164 GAGCAGAGCGGCGCGGACCTTGACCTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1223  
QY 1258 GGGACCTTCAACATCAAGCTGACCTGAGGAGACACGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 1317  
Db 1224 ATACTCATGACCATCGGCTAGCATGCTGAGGAGCTACCTGAGTGTCTATCTGAGC 1283  
QY 1318 ATGCTGGGCAATCGGCAATCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1377  
Db 1284 ATGCTGGCAATCTTGGCTTGTGGCTTCTTTGAAGTGGGTCTGCTGCTGCTGCTGCTGCTGCTG 1343  
QY 1378 ATCTTGAATGAGTCTTCCAGCAATCTCAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437  
Db 1344 TTCACTG 1403  
QY 1438 ACGTCAATGCTGCTTCCAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497  
Db 1404 TTCTCAATGAGCTCAATTTCAATTTGAGGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1463  
QY 1498 CTGAGCACCTACTGTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1557  
Db 1464 TGTGCTG 1523  
QY 1558 TATTTGTGCTGCTGAGACCAAAACAGAACCTATGAGAAATC 1602  
Db 1524 TACTTCAAGATTTCTGAGACTAAAGCGGACCTTCGATGATGATC 1568

RESULT 9  
US-09-221-107-135  
; Sequence 135, Application US/09221107  
; Patent No. 6660838  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY OF LUNG CANCER  
; FILE REFERENCE: 210121.455C2  
; CURRENT APPLICATION NUMBER: US/09/221,107  
; CURRENT FILING DATE: 1998-12-22  
; NUMBER OF SEQ ID NOS: 161  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 135  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Human  
; US-09-221-107-135

Query Match 14.0%; Score 328.6; DB 4; Length 2856;  
Best Local Similarity 53.1%; Pred. No. 2.5e-68;  
Matches 725; Conservative 0; Mismatches 634; Indels 6; Gaps 1;  
QY 238 CTCTG 297  
Db 210 CGCCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 269  
QY 298 TCGGTGGTGAATGCCCCCCCCCAGTACATCAAGGCTTTTACATGAGTCAATGGAAGA 357  
Db 270 GGAGTCATCAATGCCCCCCCCCAGAGGTGATCGAGGAGTCTTACAACAGACATGGGTCCAC 329  
QY 358 AGGCATGAGCGTCCCAATAGACCAGACACTCTGACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 417  
Db 330 CCGTATGGGAGAGCATCTGCCCCACACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 389  
QY 418 ATATTGCGCATCGTGGGAGCTTGTGGGAGCTTAATGTGAAGATGATGGAAGAGTCTT 477  
Db 390 ATCTTTCTGTTGGGCGCATGATTGGCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 449  
QY 478 GGGAGAGACACATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 537  
Db 450 GCGCGCGGAAATCAATGCTGATGAAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 509  
QY 538 GCCTGCTGCTCCAGGAGGAGCCTTTGAAATGCTCAATGTTGGGAGCGCTTCATCATGCGGC 597  
Db 510 GGTCTTCTGAAACTGGGCAAGTCTTTGAGATGCTGATCTTGGGCGCTTCATCATCGGT 569  
QY 598 ATAGATGAGGCGTCCGCTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 657  
Db 570 GTGTACTGCGGCTGACCCACAGGCTTCGTGCCCATGTATGTTGGGTGAAGTGTCAACCCACA 629  
QY 658 GAGATCGGTGGTCTCTGCGGCGAGTGAAGTGCATCTTTATCTGCTGCTGCTGCTGCTGCTGCTGCT 717  
Db 630 GCCTTCTGTTGGGCGCTTGGGCGACCTGCAACAGCTGGGCGATGCTGCTGCTGCTGCTGCTGCTGCT 689  
QY 718 GGGCAGCTTCTGGGCGCTCCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTT 777  
Db 690 GCCCAGGTGTTCGGCTGGACTTCCATCATGAGGCAACAGGACCTGTTGGCGCTGCTGCTGCTGCTG 749  
QY 778 GAGTGAATGTGGTCCCTGCGTGTCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837  
Db 750 AGCATCATCTTCACTCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 809  
QY 838 CCAAGCTTACCTGCTTGGGAGGACACACAGGAGCAAGAGTGTGAAGGCTTCCAAAG 897  
Db 810 CCGCGCTTCTGCTCATCAACCGCAACGAGGAGAACCGGCGCCAAAGTGTGCTTAAAGAG 869  
QY 898 TTCTTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGTCTTGGCTGAGAGCCACGTG 957  
Db 870 CTGCGGGAGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 929  
QY 958 CAGAGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017  
Db 930 ATGATCGGAGAGAGAGTCAACATCTGAGAGTGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 989



RESULT 6  
US-09-480-884A-135  
; Sequence 135, Application US/09480884A  
; Patent No. 6482597  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Hoeken, Nancy A.  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C6  
; CURRENT APPLICATION NUMBER: US/09/480,884A  
; CURRENT FILING DATE: 2001-08-27  
; NUMBER OF SEQ ID NOS: 330  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 135  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-480-884A-135  
Query Match 14.0%; Score 328.6; DB 4; Length 2856;  
Best Local Similarity 53.1%; Pred. No. 2.5e-68;  
Matches 725; Conservative 0; Mismatches 634; Indels 6; Gaps 1;  
QY 238 CTCCTCGTGGCTCCCTCGCGGGCGCTTCGGCTCCCTCTCTCTACGGCTCAACCTG 297  
Db 210 CGCTTCATGCTGGCTGGGAGAGCAGTCTTGGCTCCCTGCACTTTGGCTCAACACT 269  
QY 298 TCGTGTGAATGCCCGCCACCCGCTACATCAAGGCTTTTACATGAGTATGGAAGA 357  
Db 270 GGAGTCATCAATGCCCGCCAGAGGTGATCGAGAGTTCTACACACAGATGGGTCCAC 329  
QY 358 AGGATGAGCTGCAATGAGACCGACACTCTGCTCTGCTGTGCTGCTGCTGCTGCTG 417  
Db 330 CGCTATGGGAGAGCATCTCCGCCACCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 389  
QY 418 ATATTCGCATCGGTGGACTTGTGGGAGCTTAATTTGAGAGATGATTGGAAGGTTCTT 477  
Db 390 ATCTTTCTGTGGGGCATGATTGGCTCTCTCTCTGCGGCTTTTCGTTAACCGCTTT 449  
QY 478 GGGAGAGAGCACATTTCTGCGCCAAATTAATGGGTTTGCAATTTCTGCTGCAATGCTGATG 537  
Db 450 GCGCGCGGAATTCATGCTGATGATGAACCTGCTGGCTTCGTTGCTCGCGCTGCTCATG 509  
QY 538 GCCTGCTCGCTCCAGCGAGGAGCTTTGAAATGCTCATGTGGGACGCTTCATCATGGC 597  
Db 510 GCCTTCTGAACTGGGCAAGTCTTTGAGATGCTGATCCTGGGCGGCTTCATCATGGT 569  
QY 598 ATAGATGGAGCGCTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAAAG 657  
Db 570 GTGTACTCGGCTGACACAGGCTTCGTGCCATGATGTGGGTGAAGTGTACCCACA 629  
QY 658 GAGATCCGTGGCTCTCGGGGAGGTGACTGCACTCTTTATCTGCAATGGGCTGTTCAT 717  
Db 630 GCCTTCTGTGGGCTCTGGGACCTTCACAGCTGGGCATCGTGTGGCATCTCATC 689  
QY 718 GGGCAGCTTCTGGGCTGCGCGAGTGTCTGGAAAGGAGTACCTGGCCATCTGTTT 777  
Db 690 GCCAGGTGTTGGGCTGAGCTTCCATCATGGGCAACAAGGAGCTGTGGCCCTGCTGCTG 749  
QY 778 GGAGTGAATGTGGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837  
Db 750 AGCATCATCTTCATCCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 809  
QY 838 CACGCTACTGCTCTGTGAGAGGACACACGAGGAGCTGTGAAGGCTTCCAAAGC 897  
Db 810 CCGCGCTTCTGCTCTCAACCGCACGAGGAGAACCGGGCCAAAGTGTGCTAAAGAG 869  
QY 898 TTCTTGGTAAAGCAAGCTTTCCCAAGAGGTAGAGGCTCTGGCTGAGACCACTG 957

870 CTGCGGGGACAGCTGAGCTGACCCATGACCTGCAGGAGATGAAGGAGAGTCCGCGAG 929  
958 CAGAGGAGCATCCGCTGGTGTCCGCTGCTGAGTGTGAGAGCTCCCTAGTCCGCTGG 1017  
930 ATGATGCGGGAGAGAAGGTCAACATCTCTGAGCTGTTCGCTCCCGCCCTACCGCCAG 989  
1018 CAGGTGTCACCGTGAATGTCACCATGCTGCCTGCTACAGAGCTCTGTGGGCTCAATGCAATT 1077  
990 CCCATCCTCATCGCTGTGCTGCTGAGCTGCTCCAGAGCTGTCTGGCATCAACGCTGTC 1049  
1078 TGGTTCTATACACACAGCATCTTTGGAAGAGCTGGATCCCTCCGGCAAGATCCCATPAC 1137  
1050 TTCTATTACTCCAGAGCATCTTCGAGAAGCGGGGGTGCAGAGCCCTGTG-----TAT 1103  
1138 GTCACTTTGAGTACAGGGGGCATCGAGACTTTTGGCTGCCGCTCTCTCTGTTTGGTCAATT 1197  
1104 GCCACCATTTGGCTCCGGTATCGTCAACACGCGCTTCACTGCTGCTGCTGCTGCTGTTGGTG 1163  
1198 GAGCACCTGGAGACGAGACCCCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTCTTT 1257  
1164 GAGCGAGAGCGCGCGGACCCCTGCACCTCATAGGCTCTGCTGGCATGGCGGTTGTGCC 1223  
1258 GGGACCTTCACCATCACTGACCTGACCTGAGGACACGCGCCCTCGGCTCCCTTACCTGAGT 1317  
1224 ATACTCATGACCATCGGCTAGCACTGCTGGAGAGCTACCTCTGGATGCTCTATCTGAGC 1283  
1318 ATCGTGGGCACTTCGGCCATCATCGCTCTTTCTGCACTGGGCGCCAGGTGGCATCCCGTTC 1377  
1284 ATCGTGGCCATTTTGGCTTTTGGCTTTTGGCTTTTGGTGGGCTCTGTTGAGTGGGCTCTGGCCCATCCATGG 1343  
1378 ATCTTGACTGTGAGTCTTTCAGCAATCTTCAGCGGCGGCTGCTCTCATATTGCGAGC 1437  
1344 TTCACTCGTGGCTGAATCTTTCAGCCAGGCTCCAGCTCCAGTGCATTCGCTTTCAGCG 1403  
1438 ACCGTAACCTGCTCTCAACTTTGCTGTTGGGCTCTCTTCCCATTCATTCAGAAAGT 1497  
1404 TTCTCAACTGACCTCAAAATTTTCTGTTGGGCTGCTGCTTCCAGTATGTGAGCAACTG 1463  
1498 CTGGACACTACTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1557  
1464 TGTGCTCCCTAGCTCTTCTCATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1523  
1558 TATTTTGTGCTGCTGAGACCAAAACAGAACTATGCAAAATC 1602  
1524 TACTTCAAGTTCCTGAGACTAAAGCCGAGCTTCGATGAGATC 1568

RESULT 7  
US-09-542-615A-135  
; Sequence 135, Application US/09542615A  
; Patent No. 6518256  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Liqun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hoeken, Nancy A.  
; APPLICANT: Fanger, Gary R.  
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY  
; FILE REFERENCE: 210121.455C8  
; CURRENT APPLICATION NUMBER: US/09/542,615A  
; CURRENT FILING DATE: 2000-04-14  
; NUMBER OF SEQ ID NOS: 350  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 135  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-542-615A-135  
Query Match 14.0%; Score 328.6; DB 4; Length 2856;  
Best Local Similarity 53.1%; Pred. No. 2.5e-68;

Db 1258 TGACTGGTGGTCTTCTTCCAGCAATCTCAGCGCGGCTGCCTTCATCATTCATGCGCACCG 1317  
Qy 1442 TCAACTGGCTCTCCAACTTTGGTGGCTCTCTTCCCATTCATTCAGAAAAGTCTGG 1501  
Db 1318 TCAACTGGCTCTCCAACTTTGGTGGCTCTCTTCCCATTCATTCAGAAAAGTCTGG 1377  
Qy 1502 ACACCTACTGTTTCCCTAGTCTTTGCTCAAAATTCATTCAGAGGTCTATCTACCTGTATT 1561  
Db 1378 ACACCTACTGTTTCCCTAGTCTTTGCTCAAAATTCATTCAGAGGTCTATCTACCTGTATT 1437  
Qy 1562 TTGTCTCGCTGAGACCAAAACAGAACTATGAGAAATCAGCAGGCAATTTTCCAAA 1621  
Db 1438 TTGTCTCGCTGAGACCAAAACAGAACTATGAGAAATCAGCAGGCAATTTTCCAAA 1497  
Qy 1622 GGAACAAAGCATACCACCAAGAGAGAAAATCGACTCAGCTGTCACTGATGCT 1674  
Db 1498 GGAACAAAGCATACCACCAAGAGAGAAAATCGACTCAGCTGTCACTGATGCT 1550

# RESULT 5

US-09-643-597-135  
; Sequence 135, Application US/09643597  
; Patent No. 6426072  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Fan, Lijun  
; APPLICANT: Kalos, Michael D.  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Hosken, Nancy  
; APPLICANT: Fanger, Gary R.  
; APPLICANT: Li, Samuel X.  
; APPLICANT: Wang, Aijun  
; APPLICANT: Skeiky, Yasir A.W.  
; APPLICANT: Henderson, Robert A.  
; APPLICANT: McNeill, Patricia D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.455C11  
; CURRENT APPLICATION NUMBER: US/09/643,597  
; CURRENT FILING DATE: 2000-08-21  
; NUMBER OF SEQ ID NOS: 369  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 135  
; LENGTH: 2856  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-09-643-597-135

Query Match 14.0%; Score 328.6; DB 4; Length 2856;  
Best Local Similarity 53.1%; Pred No. 2.5e-68;  
Matches 725; Conservative 0; Mismatches 634; Indels 6; Gaps 1;

Qy 238 CTCCTCGTGGCTCCCTCGCGGCGCCTTGGCTCCTCCTCTACGGCTACAACTG 297  
Db 210 CGCCTCATGCTGGCTGTGGAGGAGAGTGTGGCTCCCTGCAGTTTGGCTACAACT 269  
Qy 298 TCGGTGGTAATGCCCCACCCCGTACATCAAGGCCCTTTTCAATGAGTCTATGGAAAGA 357  
Db 270 GGAGTCATCAATGCCCCCGAGAGGTGATCGAGGAGTTCTACACACAGACATGGGTCCAC 329  
Qy 358 AGCGATGACGTCCTATAGACCCACACACTCTGACTCTGCTCTGCTGTGACTGTGTCC 417  
Db 330 CGCTATGGGGAGAGCATCTCGCCACCGCTCACCACGCTCTGCTCCTCTCAGTGGCC 389  
Qy 418 ATATTGCCATCGGTGGAGCTTGTGGGAGCTTAAATTTGAGATGATTTGAAAGGTTCTT 477  
Db 390 ATCTTTTCTGTTGGGGGCAATGATGGCTCTCTCTCTGTGGGCTTTTCTGTTAACCGCTTT 449  
Qy 478 GGGAGGAACACACTTTGCTGGCCCAATATGAGGTTTGAATTTCTGCTGCAATTTGCTGATG 537  
Db 450 GCGCGCGGAATTCATGCTGATGATGAACCTGCTGGCTTGTGCTCCCGTGTCTCATG 509

Qy 538 GCCTGCTCGCTCCAGGACAGAGCCTTTGAAATGCTCAITTTGGGACGCTTCATCATGGGC 597  
Db 510 GGGTTCTCGAAACTGGGCAAGTCTTTGAGATGCTGATCCTGGGCGCTTCATCATCGGT 569  
Qy 598 ATAGATGGAGGGCTCCGCTCTAGTGTCTCCCATGTACTCAGTGAAGTCTCACCAAG 657  
Db 570 GTGTACTCGGGCTGACACAGGCTTCGTGCGCCATATATGTGGGTGAAGTGTCAACCA 629  
Qy 658 GAGATCGTGGCTCTCTGGGACAGGTGACTGCCATCTTTATCTGCAATGSGGTGTTCACT 717  
Db 630 GCCTTTCTGGGGGCTTGGGCACTTGCACAGCTGGGATCGTCTGGCATTCCTCATC 689  
Qy 718 GGGCAGCTTCTGGGCTCGCCGAGCTGTGGGAAAGGAGAGTACCTGGGCAATACCTGTTT 777  
Db 690 GCCCAGGTGTTGGCTCTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTGCTG 749  
Qy 778 GAGTGAATGTGGTCCCTGCGGTGCTCAGCTGTGAGCTTCCCTTTCTCCGGAAGC 837  
Db 750 AGCATCATCTTCAATCCGCGCTGCTGAGTGCATGCTGCTGCCCTTCTGCCCGAGAGT 809  
Qy 838 CCAGCTACTGCTCTTTGAGAAAGCAACAGGGAAGAGCTGTGAAAGCTTCCAAAG 897  
Db 810 CCGCGCTTCTGCTCATCAACCGCAACAGGAGAAACCGGCGCAAGAGTGTGTAAAGAG 869  
Qy 898 TTCTTGGGTAAAGCAGACGCTTTCCAGAGGTAGAGGAGTCTGCTGAGAGCCACGTG 957  
Db 870 CTGCGGGGACAGCTGACGTGACCCATGACCTCAGAGATGAGAGAGAGTGGCAG 929  
Qy 958 CAGAGGAGCATCCGCTGTGTGCTGCTGAGCTGCTGAGAGCTCCCTACGTCCTG 1017  
Db 930 ATGATGGGGAGAGAGGTCACTCCCTGGAGCTGTTCGCTCCCGCTACCGCCAG 989  
Qy 1018 CAGGTGGTCACTGATTTGTCACATGGCTGTCTACAGCTCTGTGGCTCAATGCAAT 1077  
Db 990 CCCATCCTCATCGCTGTGTGCTGCTGAGCTGTCCAGCAGCTGTCTGGCATCAACGCTG 1049  
Qy 1078 TGCTTCTATACCAAGACATCTTTGAAAGAGTGGATCCCTCGGCAAGATCCCATAC 1137  
Db 1050 TTCTATTACTCCACGAGCATCTTCGAGAGGCGGGGTGAGCAGGCTGTG-----TAT 1103  
Qy 1138 GTCACCTTGAGTACAGGGGCATCGAGACTTTGGCTGCGCTCTCTCTGTTGGTCAAT 1197  
Db 1104 GCCACATTTGGCTCCGGTATCGTCAACAGGCTTCACTGTCTGCTGCTGTTGGTG 1163  
Qy 1198 GAGCACCTGGGACGAGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCT 1257  
Db 1164 GAGGAGAGCGCGCGGACCTTGACCTCATAGGCTCGCTGGCATGGGGTGTGTC 1223  
Qy 1258 GGGACCTCAACATCAAGTACCCCTGACAGACACAGCCCTTGGTCCCTACCTGAGT 1317  
Db 1224 ATACTCATGACCATCGGCTAGCAGCTGTGAGCAGTACCCCTGGATGCTCTATCTGAGC 1283  
Qy 1318 ATGTTGGGCAATCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGTGGCATCCGTTTC 1377  
Db 1284 ATCGTGGCCATCTTTGGCTTTTGGCTTCTTTGAAAGTGGGTCTTGGCCCATCCCATGG 1343  
Qy 1378 ATCTTGACTGGTGGTCTTCCAGCAATCTCAGCGCGGCTGCTTCATCATTTGCAAGC 1437  
Db 1344 TTCTCGTGGCTGAACTTTTCCAGCGGCTGACAGTCCAGCTGCCATTCGCGTTGAGGC 1403  
Qy 1438 ACCGTCAACTGGCTCTCCAACTTTGTGGGCTCTCTCTTCCCATTCATTCAGAAAGT 1497  
Db 1404 TTCTCCACTGGACCTCAATTTTCAATTTGGGCAATGCTTCCAGTATGTGGAGCACTG 1463  
Qy 1498 CTGGACACTTACTGTTTCTGCTTGTGCTAATTTGTAATCAAGGTGCTATCTACTG 1557  
Db 1464 TGTGTCCTTACGCTTCTCATCTGCTCTGCTGCTCTGCTGCTGCTTCTTCTTCTCACT 1523  
Qy 1558 TATTTGTGCTGCTGAGACCAAAACAGAACCTATGCAAAATC 1602  
Db 1524 TACTTCAAGTTCTGAGACTAAAGCCGAGCTTCGATGAGATC 1568

```

Db          2341 CGC 2343

RESULT 4
US-09-620-312D-574
; Sequence 574, Application US/09620312D
; Patent No. 6569682
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yungqing
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; FILE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP2B
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pc_FL_genes Version 1.0
; SEQ ID NO 574
; LENGTH: 2842
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (345)..(1571)
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(2842)
; OTHER INFORMATION: n = a,t,c or g
US-09-620-312D-574

Query Match          56.2%; Score 1317.8; DB 4; Length 2842;
Best Local Similarity 89.9%; Pred. No. 1.8e-302;
Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 2;

QY      2  CGATCCACGCGTCGGGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGT 61
DB      40  CGATCCACGCGTGC-GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGT 98

QY      62  CACTTGAGACCCATGGCAGGAAACAAATAAGGAATTCGAAGAACTGGGCCTAGTTCCTCC 121
DB      99  CACTTGAGACCCATGGCAGGAAACAAATAAGGAATTCGAAGAACTGGGCCTAGTTCCTCC 158

QY     122  TCACAGATGACACAGCCACGCCGGGCTCCAGCGGCCAGGAGGACACTGCTGGAGTGTG 181
DB     159  TCACAGATGACACAGCCACGCCAGGCTCCAGGCTGGAAGAGGAGAAAGAAAGACTGCTGGAGTGTG 218

QY     182  ACCACCTGAGAGTGGGGTGCCAGGTGGGAAGAGAGAAAGACTGGTCTCTGCTGCCTCC 241
DB     219  ACCACCTGAGAGTGGGGTGCCAGGTGGGAAGAGGAGAAAGAAAGACTGCTGCTGCCTCC 278

QY     242  TCGTGGGCTCCCTCGCGGGGCGCCTTCGGGCTCCTCCCTTCTCTACGGCTACAACTGTGCG 301
DB     279  TCGTGGGCTCCCTCGCGGGGCGCCTTCGGGCTCCTCCCTTCTCTACGGCTACAACTGTGCG 338

```

	Db	1261	ACCCCTCACCATCAGCGTCAACCCTGCAGGACACAGCCCCCTGGGTCCCCTACTCTGAGTATC	1320
	Qy	1321	GTCGGCATTTCTGGCCATCATCGCTCTTTCTGCACTGGGCCAGGTGGCATCCCGTTCA TC	1380
	Db	1321	GTCGGCATTTCTGGCCATCATCGCTCTTTCTGCACTGGGCCAGGTGGCATCCCGTTCA TC	1380
	Qy	1381	TTGACTGTGTAGTTCTTCAGAAATCTCAGGGCCGGTGCCCTTCATCATTCAGAGCAC C	1440
	Db	1381	TTGACTGTGTAGTTCTTCAGAAATCTCAGGGCCGGTGCCCTTCATCATTCAGAGCAC C	1440
	Qy	1441	GTCAA CTGGCTCTCCA ACTTTGCTGTGGGCTCTCTTCCCAATTCATT CAGAAAAGTCTG	1500
	Db	1441	GTCAA CTGGCTCTCCA ACTTTGCTGTGGGCTCTCTTCCCAATTCATT CAGAAAAGTCTG	1500
	Qy	1501	GACAC TACTGTTTTCTTAGTCTTTGCTACAATTTTGTA TCACAGGTGCTATCTACCTGTAT	1560
	Db	1501	GACAC TACTGTTTTCTTAGTCTTTGCTACAATTTTGTA TCACAGGTGCTATCTACCTGTAT	1560
	Qy	1561	TTTGTGCTCCCTGAGACCAAAAAACAGAACCTATGCAGAAATCAGCCAGGCATTTTCCAAA	1620
	Db	1561	TTTGTGCTCCCTGAGACCAAAAAACAGAACCTATGCAGAAATCAGCCAGGCATTTTCCAAA	1620
	Qy	1621	AGGAACAAAGCATACCCACGAGAAGAAAATTCGACTCAGCTGTCTACTGTATGCTCTGCT	1680
	Db	1621	AGGAACAAAGCATACCCACGAGAAGAAAATTCGACTCAGCTGTCTACTGTATGCTCTGCT	1680
	Qy	1681	TCTTCTCTTTTCACTACTCCGAAATACAGGCTGGATTCAGCTCAGCTGTCTACTGTATGCTCTGCT	1740
	Db	1681	TCTTCTCTTTTCACTACTCCGAAATACAGGCTGGATTCAGCTCAGCTGTCTACTGTATGCTCTGCT	1740
	Qy	1741	ACCAAAAAAGAACACCCATTGTAAACGGTCATGTGGTATTTCTCAACCTCGAATGACCT	1800
	Db	1741	ACCAAAAAAGAACACCCATTGTAAACGGTCATGTGGTATTTCTCAACCTCGAATGACCT	1800
	Qy	1801	TCCCTCTATCTTCTCTCTGGAGAACACCAAGTCATGATGTGCAGACAAGAGCTTGGATTT	1860
	Db	1801	TCCCTCTATCTTCTCTCTGGAGAACACCAAGTCATGATGTGCAGACAAGAGCTTGGATTT	1860
	Qy	1861	TGGAGACATGGGTTTGAATCCAGTCATTCATTTCTTTTATTCAGCAAAATATTTAA CAAGT	1920
	Db	1861	TGGAGACATGGGTTTGAATCCAGTCATTCATTTCTTTTATTCAGCAAAATATTTAA CAAGT	1920
	Qy	1921	ACTGACATGTCCCATATGTTGTTTTACCCACCTGGTTTATA CAATGGGAGGGAGAGAGAG	1980
	Db	1921	ACTGACATGTCCCATATGTTGTTTTACCCACCTGGTTTATA CAATGGGAGGGAGAGAGAG	1980
	Qy	1981	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTAGGCTGTGCACCGTGGCTC	2040
	Db	1981	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTGAAGTCTAGGCTGTGCACCGTGGCTC	2040
	Qy	2041	ACGCTGTAACTCCAGCACTTTGGGAGGCCGAGGTGGGTGGATTCGTTGAGT CAGGAGATT	2100
	Db	2041	ACGCTGTAACTCCAGCACTTTGGGAGGCCGAGGTGGGTGGATTCGTTGAGT CAGGAGATT	2100
	Qy	2101	GAGACCATCTGGCTTAACATGGTGAAACTCCCTCTCTACTTAA AATAACAAAAATTTAGCT	2160
	Db	2101	GAGACCATCTGGCTTAACATGGTGAAACTCCCTCTCTACTTAA AATAACAAAAATTTAGCT	2160
	Qy	2161	GAGCATGGTGGCGGGCGCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCGAGGAGATGGC	2220
	Db	2161	GAGCATGGTGGCGGGCGCTGTAGTCCAGCTACTTTGGGAGGCTGAGGCGAGGAGATGGC	2220
	Qy	2221	GTGAACCCAGGAGGGGAGCTTCAGGTAGCCAGATCACACCA CACACTCCAGGCTGG	2280
	Db	2221	GTGAACCCAGGAGGGGAGCTTCAGGTAGCCAGATCACACCA CACACTCCAGGCTGG	2280
	Qy	2281	GTGACAGAGCCAGACTCCGCTCTCAAAAAA AAAAAAAAAAAAAAAAAAGGGCGGC	2340
	Db	2281	GTGACAGAGCCAGACTCCGCTCTCAAAAAA AAAAAAAAAAAAAAAAAAGGGCGGC	2340
	Qy	2341	CGC 2343	

2341 CGC 2343



```

Db      2281 GTGACAGCCAGACTCCGTCCTCAAAAAAAAAAAAAAAAAAGGCGGC 2343
QY      2341 CGC 2343
Db      2341 CGC 2343

RESULT 3
US-09-610-417-1
; Sequence 1, Application US/09610417
; Patent No. 6346374
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/610,417
; FILING DATE: 05-Jul-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/299,549
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELE: 200154
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2343 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 73...1761
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-610-417-1

```

	Query Match	100.0%;	Score 2343;	DB 4;	Length 2343;
	Best Local Similarity	100.0%;	Pred. No. 0;		
	Matches 2343;	Conservative	0;	Mismatches	0; Indels
					0; Gaps
Qy	1	TCGACCCACGCGTCGGGCGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG	60		
Db	1	TCGACCCACGCGTCGGGCGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG	60		
Qy	61	TCACTGAGACCCCATGGCAAGAACAAAAATAGGAATTCCAAGGAACCTGGGCTAGTTCCC	120		
Db	61	TCACTGAGACCCCATGGCAAGAACAAAAATAGGAATTCCAAGGAACCTGGGCTAGTTCCC	120		
Qy	121	CTCAGATGACACACAGCCACGCGGGCGCTCCAGGGCGAAGGAGGGCACTGCTGGAGTGT	180		
Db	121	CTGACGATGACACACAGCCACGCGGGCGCTCCAGGGCGAAGGAGGGCACTGCTGGAGTGT	180		





Db 1 TCGACCCACGCGTCCGGCTTGGCAGAGTCTGGGTCCCTGACTGAGCCATCAGCTGGG 60  
Qy 61 TCACTGAGACCCATGCGAAGGAAACAAATAAGAAATTCAGAGAACTGGGCTAGTTCCC 120  
Db 61 TCACTGAGACCCATGCGAAGGAAACAAATAAGAAATTCAGAGAACTGGGCTAGTTCCC 120  
Qy 121 CTCACAGATGACACACGACCCGCGCTCCAGGCGCAGAGGAGGCACTGCTGAGTGT 180  
Db 121 CTCACAGATGACACACGACCCGCGCTCCAGGCGCAGAGGAGGCACTGCTGAGTGT 180  
Qy 181 GACCACTGAGGAGTGGGTGCGCAGGTGGAAGGAGAAAGAACTGGTCTGCTGCTC 240  
Db 181 GACCACTGAGGAGTGGGTGCGCAGGTGGAAGGAGAAAGAACTGGTCTGCTGCTC 240  
Qy 241 CTGCTGGCTCCCTCGCGGCGCTTGGGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
Db 241 CTGCTGGCTCCCTCGCGGCGCTTGGGCTCCTCTCTCTCTCTCTCTCTCTCTCTCT 300  
Qy 301 GTGGTGAATGCCCCACCCGCTACATCAAGGCGCTTACATGAGTCAATGGAAGAGG 360  
Db 301 GTGGTGAATGCCCCACCCGCTACATCAAGGCGCTTACATGAGTCAATGGAAGAGG 360  
Qy 361 CATGAGCGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTCTGCTCTGCTC 420  
Db 361 CATGAGCGTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTCTGCTCTGCTC 420  
Qy 421 TTGCGCATCGGTGGAAGTGTGGGAGCTTAATTTGGAAGATGATGGAAGAGTCTTGGG 480  
Db 421 TTGCGCATCGGTGGAAGTGTGGGAGCTTAATTTGGAAGATGATGGAAGAGTCTTGGG 480  
Qy 481 AGGAAGCACACTTGTGCGGCAATATGGTGTGCAATTTCTGCTGCTCTGCTCTGCTC 540  
Db 481 AGGAAGCACACTTGTGCGGCAATATGGTGTGCAATTTCTGCTGCTCTGCTCTGCTC 540  
Qy 541 TGTCTGCTCCAGGAGGAGGCTTTGAAATGCTCATTTGTTGGGACGCTTCATCATGGGATA 600  
Db 541 TGTCTGCTCCAGGAGGAGGCTTTGAAATGCTCATTTGTTGGGACGCTTCATCATGGGATA 600  
Qy 601 GATGAGAGGCTCGCCCTCAGTGTGCTCCCAAGTACCTCAGTGAATCTACCCAGAG 660  
Db 601 GATGAGAGGCTCGCCCTCAGTGTGCTCCCAAGTACCTCAGTGAATCTACCCAGAG 660  
Qy 661 ATCCGTGCTCTCTGGGAGGAGTCTGCTCACTTTATCTCATTTGGGCTTCACTGGG 720  
Db 661 ATCCGTGCTCTCTGGGAGGAGTCTGCTCACTTTATCTCATTTGGGCTTCACTGGG 720  
Qy 721 CAGCTTCTGGGCTCGCCGAGTCTGGGAAAGGAGTACCTGGCCATACCTGTTTGA 780  
Db 721 CAGCTTCTGGGCTCGCCGAGTCTGGGAAAGGAGTACCTGGCCATACCTGTTTGA 780  
Qy 781 GTGATTTGGTCCCTCGGCTTCCAGTGTGAGCTTCCCTTCTCCGCGAGAGCCCA 840  
Db 781 GTGATTTGGTCCCTCGGCTTCCAGTGTGAGCTTCCCTTCTCCGCGAGAGCCCA 840  
Qy 841 CGCTACCTCTCTTGGAGAGCACACAGAGGAGAGTGTGAAAGCTTCCAAACGCTC 900  
Db 841 CGCTACCTCTCTTGGAGAGCACACAGAGGAGAGTGTGAAAGCTTCCAAACGCTC 900  
Qy 901 TTGGTAAAGCAGAGTGTTCCTCAAGAGGTAGAGAGTCTGGGCTGAGAGCCAGTGCAG 960  
Db 901 TTGGTAAAGCAGAGTGTTCCTCAAGAGGTAGAGAGTCTGGGCTGAGAGCCAGTGCAG 960  
Qy 961 AGAGCATCCGCTGGTGTCTGTGAGTGTGAGAGTCTCCCTACGTCGCTGGCAG 1020  
Db 961 AGAGCATCCGCTGGTGTCTGTGAGTGTGAGAGTCTCCCTACGTCGCTGGCAG 1020  
Qy 1021 GTGGTCAACGATGTCACCAATGCGCTTACAGCTCTGTGGCTCAATGCAATTTGG 1080  
Db 1021 GTGGTCAACGATGTCACCAATGCGCTTACAGCTCTGTGGCTCAATGCAATTTGG 1080  
Qy 1081 TTCTATACCAACAGCATCTTTGAAAGCTGGGATCCCTCCGCGAAAGATCCCATACCTC 1140

Db 1081 TTCTATACCAACAGCATCTTTGGAAAGCTGGATCCCTCCGCGAAAGATCCCATACGTC 1140  
Qy 1141 ACCTTGAATACAGGGGATCGAGATTTGGCTGCCGCTCTCTCTGTTTGGTCAATGAG 1200  
Db 1141 ACCTTGAATACAGGGGATCGAGATTTGGCTGCCGCTCTCTCTGTTTGGTCAATGAG 1200  
Qy 1201 CACCTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTTGGG 1260  
Db 1201 CACCTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTTTGGG 1260  
Qy 1261 ACCCTCACCATCAGCCTGACCTTGACAGCACCCCTCCCTGGGCTCCCTACTCTGATATC 1320  
Db 1261 ACCCTCACCATCAGCCTGACCTTGACAGCACCCCTCCCTGGGCTCCCTACTCTGATATC 1320  
Qy 1321 GTGGGATTTCTGGCCATCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
Db 1321 GTGGGATTTCTGGCCATCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1380  
Qy 1381 TTGACTGTGTGAGTCTCTCCAGCAATCTCAGCGCGGCTGCTCTCATCATTTGACAGGAC 1440  
Db 1381 TTGACTGTGTGAGTCTCTCCAGCAATCTCAGCGCGGCTGCTCTCATCATTTGACAGGAC 1440  
Qy 1441 GTCAACTGGCTCTCCAACTTTGCTGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
Db 1441 GTCAACTGGCTCTCCAACTTTGCTGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1500  
Qy 1501 GACACTACTGTTTCTCTAGTCTTTGCTATCAATTTGATATCAAGTGTCTATCTACTGTAT 1560  
Db 1501 GACACTACTGTTTCTCTAGTCTTTGCTATCAATTTGATATCAAGTGTCTATCTACTGTAT 1560  
Qy 1561 TTTTGTGCTGCTGAGACCAAAACAGACCTATCGAGAAATCAGCAGGCAATTTTCCAAA 1620  
Db 1561 TTTTGTGCTGCTGAGACCAAAACAGACCTATCGAGAAATCAGCAGGCAATTTTCCAAA 1620  
Qy 1621 AGGAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGCTGCTGCTGCTGCTGCT 1680  
Db 1621 AGGAACAAAGCATACCCACAGAGAGAGAAATCGACTCAGCTGCTGCTGCTGCTGCTGCT 1680  
Qy 1681 TTTTCTCTCTTCACTCTCCGAAATACAGCTGCTGAGTTCAGCTGCTGCTGCTGCTGCTGCT 1740  
Db 1681 TTTTCTCTCTTCACTCTCCGAAATACAGCTGCTGAGTTCAGCTGCTGCTGCTGCTGCTGCT 1740  
Qy 1741 ACCAAAAAGAACACCCATTTGAAAGCTCATGTTGTTATTTCTTCAACCTGGAATGACCT 1800  
Db 1741 ACCAAAAAGAACACCCATTTGAAAGCTCATGTTGTTATTTCTTCAACCTTGAATGACCT 1800  
Qy 1801 TCCCTTATCTTCTTCTCTCGAGAACACCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCT 1860  
Db 1801 TCCCTTATCTTCTTCTCTCGAGAACACCAAGTCAATGCTGCTGCTGCTGCTGCTGCTGCT 1860  
Qy 1861 TGGAGACATGGGTTTGAATTCAGTCAATCTCTTTTATTTTATTCAGCAATATTTTAAAGT 1920  
Db 1861 TGGAGACATGGGTTTGAATTCAGTCAATCTCTTTTATTTTATTCAGCAATATTTTAAAGT 1920  
Qy 1921 ACTGACATGCTCCCATATGTTTATCCCTGCTTATACAAATGGGAGGAGAGAGAG 1980  
Db 1921 ACTGACATGCTCCCATATGTTTATCCCTGCTTATACAAATGGGAGGAGAGAGAGAG 1980  
Qy 1981 AGAGAGAGAGAGAGAGTGTATTTCTAAAGCTTTGAAGTCTAGGCTGTGCAAGTGGCTC 2040  
Db 1981 AGAGAGAGAGAGAGAGTGTATTTCTAAAGCTTTGAAGTCTAGGCTGTGCAAGTGGCTC 2040  
Qy 2041 AGGCTGTAAATCCAGACACTTTGGAGGCGAGTGGTGGATCGTGGTCAAGGATTT 2100  
Db 2041 AGGCTGTAAATCCAGACACTTTGGAGGCGAGTGGTGGATCGTGGTCAAGGATTT 2100  
Qy 2101 GAGACATCTCTGGCTTACATGCTGAACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160  
Db 2101 GAGACATCTCTGGCTTACATGCTGAACTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2160  
Qy 2161 GAGCATGTGTGGGCGCTGTAGTCTCCAGCTACTTTGGAGGCTGTAGGAGGAGATGGC 2220  
Db 2161 GAGCATGTGTGGGCGCTGTAGTCTCCAGCTACTTTGGAGGCTGTAGGAGGAGATGGC 2220

(without alignment)  
7387.784 Million cell updates/sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2343	100.0	2343	2	US-09-031-392-1	Sequence 1, Appli
2	2343	100.0	2343	3	US-09-299-548-1	Sequence 1, Appli
3	2343	100.0	2343	4	US-09-610-417-1	Sequence 1, Appli
4	1317.8	56.2	2842	4	US-09-620-312D-574	Sequence 574, App
5	328.6	14.0	2856	4	US-09-643-597-135	Sequence 135, App
6	328.6	14.0	2856	4	US-09-480-884A-135	Sequence 135, App
7	328.6	14.0	2856	4	US-09-542-615A-135	Sequence 135, App
8	328.6	14.0	2856	4	US-09-608-621B-135	Sequence 135, App
9	328.6	14.0	2856	4	US-09-221-107-135	Sequence 135, App
C 10	276.6	11.8	12619	4	US-09-616-289-49	Sequence 49, Appli
C 11	275.2	11.7	320	1	US-08-629-938-5	Sequence 5, Appli
C 12	275.2	11.7	320	1	US-08-759-873-5	Sequence 5, Appli
13	273.2	11.7	246240	2	US-08-724-394A-20	Sequence 20, Appli
14	273.2	11.7	246240	2	US-08-724-394A-21	Sequence 21, Appli
15	273.2	11.7	246240	2	US-08-724-394A-22	Sequence 22, Appli
16	270.4	11.5	193303	4	US-09-497-855A-37	Sequence 37, Appli
17	270.4	11.5	193303	4	US-09-497-855A-44	Sequence 44, Appli
C 18	269	11.5	8220	4	US-09-797-908-3	Sequence 3, Appli
C 19	268.6	11.5	14636	3	US-09-173-914-6	Sequence 6, Appli
20	268.4	11.5	3742	1	US-08-694-915-5	Sequence 5, Appli
21	266.4	11.4	116592	4	US-09-818-512-3	Sequence 3, Appli
C 22	266	11.4	62804	4	US-09-800-960-3	Sequence 3, Appli
23	266	11.4	62804	4	US-10-096-960-3	Sequence 3, Appli
24	264.4	11.3	2592	4	US-09-591-025-8	Sequence 8, Appli
25	264.4	11.3	2592	4	US-09-894-927B-8	Sequence 8, Appli
26	263	11.2	22481	4	US-08-367-841A-43	Sequence 43, Appli
27	263	11.2	22481	5	PCR-US95-07201-43	Sequence 43, Appli

1 TCGACCCACGCGTCCGGCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG 60

QY 397 CTCTGGTCTGTGACGTGTCTCCATATTCGCCATCGGTGACCTGTGGGACGTTAATTGTG 456  
Db |||||  
QY 397 CTCTGGTCTGTGACGTGTCTCCATATTCGCCATCGGTGACCTGTGGGACGTTAATTGTG 396  
Db |||||  
QY 457 AAGATGATGAAAGGTTCTTGGGAGGAGACACACTTTCTGGCCATATATGGGTTTGCA 516  
Db |||||  
QY 397 AAGATGATGAAAGGTTCTTGGGAGGAGACACACTTTCTGGCCATATATGGGTTTGCA 456  
Db |||||  
QY 517 ATTTCTGTCTGATTCCTGATGCTCGTCTCCAGGACGAGCCTTTGAAATGCTCAT 576  
Db |||||  
QY 457 ATTTCTGTCTGATTCCTGATGCTCGTCTCCAGGACGAGCCTTTGAAATGCTCAT 516  
Db |||||  
QY 577 GTGGACGCTTCATCATGGGCATAGATGAGGCGTTCGCCCTCAGTGTGCTCCCCATGTAC 636  
Db |||||  
QY 517 GTGGACGCTTCATCATGGGCATAGATGAGGCGTTCGCCCTCAGTGTGCTCCCCATGTAC 576  
Db |||||  
QY 637 CTCAGTGAATCTACCCAGAGATCCGTGGCTCTCTGGGCGAGTGAATCTT 696  
Db |||||  
QY 577 CTCAGTGAATCTACCCAGAGATCCGTGGCTCTCTGGGCGAGTGAATCTT 636  
Db |||||  
QY 697 ATCTGCATGGGCTGTTCACTGGGAGCTTCTGGGCTGCGCGAGCTCTGGGAAAGGAG 756  
Db |||||  
QY 637 ATCTGCATGGGCTGTTCACTGGGAGCTTCTGGGCTGCGCGAGCTCTGGGAAAGGAG 696  
Db |||||  
QY 757 AGTACTGGCCATACCTGTTTGGAGTATTGTGGTCCCTGCCCTTGTCCAGCTGTGAGC 816  
Db |||||  
QY 697 AGTACTGGCCATACCTGTTTGGAGTATTGTGGTCCCTGCCCTTGTCCAGCTGTGAGC 756  
Db |||||  
QY 817 CTTCCCTTCTCCCGAGCAGCCACGCTACCTGCTCTTGGAGAGCAACAAGGCAAGA 876  
Db |||||  
QY 757 CTTCCCTTCTCCCGAGCAGCCACGCTACCTGCTCTTGGAGAGCAACAAGGCAAGA 816  
Db |||||  
QY 877 GCTGTGAAAGCCTTCCAAAGTCTTGGGTAAAGCAGACGTTTCCAAAGAGGTAGAGAG 936  
Db |||||  
QY 817 GCTGTGAAAGCCTTCCAAAGTCTTGGGTAAAGCAGACGTTTCCAAAGAGGTAGAGAG 876  
Db |||||  
QY 937 GTCTGGCTGAGAGCAGCGTGAGAGAGCATCCGCTGGTGTCCGTCTGGAGCTGCTG 996  
Db |||||  
QY 877 GTCTGGCTGAGAGCAGCGTGAGAGAGCATCCGCTGGTGTCCGTCTGGAGCTGCTG 936  
Db |||||  
QY 997 AGAGTCCCTAGTCCGCTGGGAGTGTACCGTGTATTCACCATGGCCTGTACCAG 1056  
Db |||||  
QY 937 AGAGTCCCTAGTCCGCTGGGAGTGTACCGTGTATTCACCATGGCCTGTACCAG 996  
Db |||||  
QY 1057 CTCTGTGGCCTCAATGCAATTTGGTTCTATACCAACAGCATCTTTGGAAAAGCTGGGATC 1116  
Db |||||  
QY 997 CTCTGTGGCCTCAATGCAATTTGGTTCTATACCAACAGCATCTTTGGAAAAGCTGGGATC 1056  
Db |||||  
QY 1117 CCTCCGGCAAGATCCCAATACATCACTTGTAGTAAGGGGGCATCGAGACTTTGGCTGCC 1176  
Db |||||  
QY 1057 CCTCTGGCAAGATCCCAATACATCACTTGTAGTAAGGGGGCATCGAGACTTTGGCTGCC 1116  
Db |||||  
QY 1177 GTCTTCTCTGGTTTGGTCAATTGAGACCTGGGACGAGACCCCTCTCATTTGGTGGCTTT 1236  
Db |||||  
QY 1117 GTCTTCTCTGGTTTGGTCAATTGAGACCTGGGACGAGACCCCTCTCATTTGGTGGCTTT 1176  
Db |||||  
QY 1237 GGGCTCATGGCCTCTTCTTTGGGACCTCACCATCAGCTGACCTGCAGGACCAAGCC 1296  
Db |||||  
QY 1177 GGGCTCATGGCCTCTTCTTTGGGACCTCACCATCAGCTGACCTGCAGGACCAAGCC 1236  
Db |||||  
QY 1297 CCTGGGCTCCCTACCTGAGTATCGTGGGCAATCTGGCCATCATGCCCTCTTTCTGCAGT 1356  
Db |||||  
QY 1237 CCTGGGCTCCCTACCTGAGTATCGTGGGCAATCTGGCCATCATGCCCTCTTTCTGCAGT 1296  
Db |||||  
QY 1357 GGGCCAGGTGGCATCCC 1373  
Db |||||  
QY 1297 GGGCCAGGTGTTTCCC 1313  
Db |||||

Search completed: April 8, 2004, 11:57:02  
Job time : 949 secs

QY	1081	TTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGGCAAGATCCCATACGTC	1140
Db	632	TTCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGGCAAGATCCCATACGTC	691
QY	1141	ACCTTTGAGTACAGGGGGGATCGAGACTTTGGCTGCGCTTCTCTCTGTTGGTCAATTGAG	1200
Db	692	ACCTTTGAGTACAGGGGGGATCGAGACTTTGGCTGCGCTTCTCTCTGTTGGTCAATTGAG	751
QY	1201	GACCTGGGACGAGACCCCTCTCTCAATGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG	1260
Db	752	GACCTGGGACGAGACCCCTCTCTCAATGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG	811
QY	1261	ACCTCCACCATCACGCTGACCCCTGACGACACGACGCGCCCTCGGCTCCCTACCTGAGTATC	1320
Db	812	ACCTCCACCATCACGCGGACCTGACGACACGACGCGCCCTCGGCTCCCTACCTGAGTATC	871
QY	1321	GTGGGCAATTCGGCCATCATGCGCTCTTTCTGAGTGGGCGGAGTGCGATCCCGTTCATC	1380
Db	872	GTGGGCAATTCGGCCATCATGCGCTCTTTCTGAGTGGGCGGAGTGCGATCCCGTTCATC	931
QY	1381	TTGACTGGTGGAGTCTTCTCAGCAATCTCAGCGGCGGCTGCTTTCATCATTCGAGGCACC	1440
Db	932	TTGACTGGTGGAGTCTTCTCAGCAATCTCAGCGGCGGCTGCTTTCATCATTCGAGGCACC	991
QY	1441	GTCAACTGGCTCTCAATTTGCTGTGGTCTCTTCCATTCATTCAGAAAGCTG	1500
Db	992	GTCAACTGGCTCTCAATTTGCTGTGGTCTCTTCCATTCATTCAGAAAGCTG	1051
QY	1501	GACACTGCTGTTTCTAGTCTTTGCTACAAATTTGATACAGGTGCTATCTACCTGTAT	1560
Db	1052	GACACTGCTGTTTCTAGTCTTTGCTACAAATTTGATACAGGTGCTATCTACCTGTAT	1111
QY	1561	TTTGTGCTGCTGAGCAAAAACAGAAACCTATCGAATAATCAGCCAGGATTTTCCAAA	1620
Db	1112	TTTGTGCTGCTGAGCAAAAACAGAAACCTATCGAATAATCAGCCAGGATTTTCCAAA	1171
QY	1621	AGGACAAGCATACCCACGAGAGAGAAATGACTGCTGCTACCTGATGCT	1674
Db	1172	AGGACAAGCATACCCACGAGAGAGAAATGACTGCTGCTACCTGATGCT	1225
RESULT 15			
ID	AAD09557		
AC	AAD09557 standard; cDNA; 2216 BP.		
XX	AAD09557;		
XX	10-SEP-2001 (first entry)		
XX	Human transporter and ion channel-6 (TRICH-6) cDNA.		
XX	Human; transporter and ion channel-6; TRICH-6; vaccine; cystic fibrosis;		
KW	gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;		
KW	hypertension; angina; neurological disorder; asthma; bipolar disorder;		
KW	dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;		
KW	Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;		
KW	Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;		
KW	demyelinating disease; mental disorder; Schizophrenia; polymyositis;		
KW	muscle disorder; cardiomyopathy; cataract; myocardiitis; Grave's disease;		
KW	dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;		
KW	rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;		
KW	sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;		
KW	scleroderma; pulmonary artery stenosis; neotropic; Addison's disease;		
KW	malabsorption syndrome; hypercholesterolaemia; cancer; ss.		
XX	Homo sapiens.		
OS			
XX	Location/Qualifiers		
FT	100..1350		
FT	/*tag= a		
FT	/product= "Human TRICH-6 protein"		
FT	100..210		
TT	sig_peptide		

```

PT /tag= a
PT /product= "Human TRICH-6 protein"
PT 100. .210
PT sig_peptide

```

QY 1380 CTTGACTGGTGGTCTTTCCAGCAATCTCAGCGCGCGGCTGCTTCATCATTCAGGCAC 1439  
Db 1141 CTTGACTGGTGGTCTTTCCAGCAATCTCAGCGCGCGGCTGCTTCATCATTCAGGCAC 1200  
QY 1440 CTTCAACTGGCTCTCAACTTTGCTGTTGGGCTCCTCTCCCATTCATTCAGAAAGTCT 1499  
Db 1201 CTTCAACTGGCTCTCAACTTTGCTGTTGGGCTCCTCTCCCATTCATTCAGAAAGTCT 1260  
QY 1500 GGACACCTACTGTTTCTAGTCTTTGCTACAAATTTGTATACAGTGTATCTACCTGTA 1559  
Db 1261 GGACACCTACTGTTTCTAGTCTTTGCTACAAATTTGTATACAGTGTATCTACCTGTA 1320  
QY 1560 TTTTGTGCTGCTGACACCAAAAAAGAACCTATGAGAAATCAGCAGGCAATTTCCAA 1619  
Db 1321 TTTTGTGCTGCTGACACCAAAAAAGAACCTATGAGAAATCAGCAGGCAATTTCCAA 1380  
QY 1620 AGGACAAAGCATCCCAACCAAGAGAGAAATCGACTCAGTGTCACTGATGCT 1674  
Db 1381 AAGGAACAAAGCATCCCAACCAAGAGAGAAATCGACTCAGTGTCACTGATGCT 1435

RESULT 14  
ID ADC26264 standard; DNA; 1267 BP.  
XX AC ADC26264;  
XX DT 18-DEC-2003 (first entry)  
XX DE Human NOV17h DNA.  
XX KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;  
KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;  
KW transgenic; human; gene; ds.  
XX OS Homo sapiens.  
XX PN WO2003004687-A2.  
XX PD 16-JAN-2003.  
XX PF 03-JUL-2002; 2002WO-US021361.  
XX PR 05-JUL-2001; 2001US-0303046P.  
XX PR 09-JUL-2001; 2001US-0303828P.  
XX PR 09-JUL-2001; 2001US-0304016P.  
XX PR 11-JUL-2001; 2001US-0304502P.  
XX PR 13-JUL-2001; 2001US-0305262P.  
XX PR 16-JUL-2001; 2001US-0305673P.  
XX PR 17-JUL-2001; 2001US-0305085P.  
XX PR 24-JUL-2001; 2001US-0307536P.  
XX PR 27-JUL-2001; 2001US-0308228P.  
XX PR 30-JUL-2001; 2001US-0308877P.  
XX PR 01-AUG-2001; 2001US-0309255P.  
XX PR 17-AUG-2001; 2001US-0311328P.  
XX PR 12-SEP-2001; 2001US-0318711P.  
XX PR 19-SEP-2001; 2001US-0323380P.  
XX PR 21-SEP-2001; 2001US-0323969P.  
XX PR 04-JAN-2002; 2002US-0345022P.  
XX PR 04-JAN-2002; 2002US-0345038P.  
XX PR 28-FEB-2002; 2002US-0361172P.  
XX PR 01-MAR-2002; 2002US-0360814P.  
XX PR 01-MAR-2002; 2002US-0360830P.  
XX PR 01-MAR-2002; 2002US-0361133P.  
XX PR 01-MAR-2002; 2002US-0361147P.  
XX PR 05-MAR-2002; 2002US-0361677P.  
XX PR 02-APR-2002; 2002US-0363637P.  
XX PR 12-APR-2002; 2002US-0372326P.  
XX PR 16-APR-2002; 2002US-0372990P.  
XX PR 19-APR-2002; 2002US-0373881P.  
XX PR 19-APR-2002; 2002US-0373921P.  
XX PR 02-JUL-2002; 2002US-00188186.

PA (CURA-) CURAGEN CORP.  
XX Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ,  
PI Catterson E, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Gorman L;  
PI Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE,  
PI Padigaru M, Patturajan M, Pena CEA, Rastelli L, Shency S;  
PI Shimkets RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ;  
PI Vernet CAM, Voss EZ, Zhong M;  
XX WPI; 2003-221607/21.  
DR P-PSDB; ADC26265.  
XX  
PT New isolated NOVX polypeptide, useful for determining the presence of, or  
PT predisposition to a disease associated with altered levels of expression  
PT of the polypeptide, and for treating or preventing cancer.  
XX  
PS Claim 20; SEQ ID NO 89; 478pp; English.  
XX  
CC The invention relates to a novel isolated NOV polypeptide. The  
CC polypeptide of the invention demonstrates cytosolic activity and may be  
CC used for determining the presence of, or predisposition to a disease  
CC associated with altered levels of expression of the polypeptide,  
CC including metabolic disorders, immune disorders, neurodegenerative  
CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases  
CC and cancer. The polypeptide may also be utilised during gene therapy  
CC procedures, vaccine development and transgenic animal production. The  
CC current sequence is that of the human NOV DNA of the invention.  
XX  
SQ Sequence 1267 BP; 251 A; 368 C; 339 G; 309 T; 0 U; 0 Other;

Query Match 50.7%; Score 1187.6; DB 9; Length 1267;  
Best Local Similarity 99.7%; Pred. No. 1.7e-265;  
Matches 1190; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 481 AGGAAGCACACTTTGCTGGCCAAATATGGTTTGGCAATTTCTGCTGCATTCGTATGGCC 540  
Db 32 AGGAAGCACACTTTGCTGGCCAAATATGGTTTGGCAATTTCTGCTGCATTCGTATGGCC 91  
QY 541 TGCTCGCTCCAGGACGAGCCTTTGAAATGCTCATTTGTGGAGCGCTTCATCATGGGCATA 600  
Db 92 TGCTCGCTCCAGGACGAGCCTTTGAAATGCTCATTTGTGGAGCGCTTCATCATGGGCATA 151  
QY 601 GATGGAGCGCTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCAAGGAG 660  
Db 152 GATGGAGCGCTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCAAGGAG 211  
QY 661 ATCCGTGGCTCTCTGGGGCAGGTGACGCCATCTTTATCTGCAATTTGGCGTTCACCTGGG 720  
Db 212 ATCCGTGGCTCTCTGGGGCAGGTGACGCCATCTTTATCTGCAATTTGGCGTTCACCTGGG 271  
QY 721 CAGCTTCTGGGCTGCCCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGA 780  
Db 272 CAGCTTCTGGGCTGCCCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGA 331  
QY 781 GTGATTGTGCTCGCTGCGCTTGTCCAGCTGCTGAGCCTTCCCTTTCTCCCGGACAGCCCA 840  
Db 332 GTGATTGTGCTCGCTGCGCTTGTCCAGCTGCTGAGCCTTCCCTTTCTCCCGGACAGCCCA 391  
QY 841 CGCTACCTGCTCTTGGAGAAGCACACAGGCGCAAGAGCTGTGAAAGCCTTCCAAACGTTTC 900  
Db 392 CGCTACCTGCTCTTGGAGAAGCACACAGGCGCAAGAGCTGTGAAAGCCTTCCAAACGTTTC 451  
QY 901 TTGGGTAAGACAGACGTTTCCCAGAGGTAGAGAGGTCTCGCTGAGAGCAGCTGACAG 960  
Db 452 TTGGGTAAGACAGACGTTTCCCAGAGGTAGAGAGGTCTCGCTGAGAGCAGCTGACAG 511  
QY 961 AGGAGCATCGCCTGCTGCTCGCTGCTGAGAGCTCCCTACGTCGGTGGCAG 1020  
Db 512 AGGAGCATCGCCTGCTGCTCGCTGCTGAGAGCTCCCTACGTCGGTGGCAG 571  
QY 1021 GTGGTACCGTGATTGTCAACATGGCCTGCTACCAAGCTCTGTGGCCTCAATGCAATTGG 1080  
Db 572 GTGGTACCGTGATTGTCAACATGGCCTGCTACCAAGCTCTGTGGCCTCAATGCAATTGG 631

	Query Match	52.5%;	Score 1229;	DB 9;	Length 1502;
	Best Local Similarity	88.2%;	Pred. No. 4.6e-275;		
	Matches 1425;	Conservative	0;	Mismatches 180;	Indels 180;
					Gaps 1;
Qy	60	GTCACTGAGACCCATGCGCAAGGAAACAAAATAGGAATTC	CAAGGAACCTGGGCTAGTTGCC	119	
Db	1	GTCACTGAGACCCATGCGCAGGAAACAAATAGGAATTC	CAAGGAACCTGGGCTAGTTGCC	60	
Qy	120	CCTCAAGATGACACCAAGCCACGCGGGCCTCCAGGGCCAGGAGGGA	CACTCCTGGAGTG	179	
Db	61	CCTCAAGATGACACCAAGCCACGCGGGCCTCCAGGGCCAGGAGGGA	CACTCCTGGAGTG	120	
Qy	180	TCACACACCTGAGGAGTGGGGTGCCAGGTGGAAGGAGAAAGGACT	TGGTCTGCTGCCT	239	
Db	121	TGTCACACCTGAGGAGTGGGGTGCCAGGTGGAAGGAGAAAGGACT	TGGTCTGCTGCCT	180	
Qy	240	CCTCGTGGCCTCCTTCGCGGGCGGCTTCGGGCTCTCTCTCTCTA	CGGCTCAACAACCTGTC	299	
Db	181	CCTCGTGGCCTCCTTCGCGGGCGGCTTCGGGCTCTCTCTCTCTA	CGGCTCAACAACCTGTC	240	



1065 CCTCAATGCAATTTGGTTCTTATACCAACAGCATCTTTGGAAAAGCTGGGATCCCTCCGGC 1124  
 1081 CCTCAATGCAATTTGGTTCTTATACCAACAGCATCTTTGGAAAAGCTGGGATCCCTCCGGC 1140  
 1125 AAAGATCCCATAGCTCAGCTTGGTACAGGGGGCATCGAGACTTTGGGTGCGCTCTTCTC 1184  
 1141 AAAGATCCCATAGCTCAGCTTGGTACAGGGGGCATCGAGACTTTGGGTGCGCTCTTCTC 1200  
 1185 TGGTTTGGTCAATAGCACTGGAGCGGAGACCCCTCTCTTCAATTTGGTGGCTTTGGGCTCAT 1244  
 1201 TGGTTTGGTCAATAGCACTGGAGCGGAGACCCCTCTCTTCAATTTGGTGGCTTTGGGCTCAT 1260  
 1245 GGCCCTCTTCTTTGGGACCCCTCAGCATCAGCTGAGCCCTTGCAGGACCAAGCCCTTGGGT 1304  
 1261 GGCCCTCTTCTTTGGGACCCCTCAGCATCAGCTGAGCCCTTGCAGGACCAAGCCCTTGGGT 1320  
 1305 CCCTACCTGAGTATCGTGGGCAATTTGGGCAATCATCGCTCTTCTTCTGAGTGGGCCAGG 1364  
 1321 CCCTACCTGAGTATCGTGGGCAATTTGGGCAATCATCGCTCTTCTTCTGAGTGGGCCAGG 1380  
 1365 TGGCATCCCGTTCACTTTGACTGGTGGTCTTCTTCCAGCAATCTCAGCGGCCGCTTGCCTT 1424  
 1381 TGGCATCCCGTTCACTTTGACTGGTGGTCTTCTTCCAGCAATCTCAGCGGCCGCTTGCCTT 1440  
 1425 CATCATGTGAGGACCGTCAACTGGCTCTTCCAACTTTGCTGCTGTTGGGCTCTTCTTCCCAT 1484  
 1441 CATCATGTGAGGACCGTCAACTGGCTCTTCCAACTTTGCTGCTGTTGGGCTCTTCTTCCCAT 1500  
 1485 CATTCAGAAAGCTTCGACACCTACTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1544  
 1501 CATTCAGAAAGCTTCGACACCTACTGTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1560  
 1545 TGCTATCTACCTGTATTTTGTGCTGCTGAGACCAAAACAGAACTTATGCAGAAATCAG 1604  
 1561 TGCTATCTACCTGTATTTTGTGCTGCTGAGACCAAAACAGAACTTATGCAGAAATCAG 1620  
 1605 CAGGCAATTTCCAAAGGAAACAAAGCATPACCCACAGAGAGAAATCGACTCAGCTGT 1664  
 1621 CAGGCAATTTCCAAAGGAAACAAAGCATPACCCACAGAGAGAAATCGACTCAGCTGT 1680  
 1665 CACTGATGCT 1674  
 1681 CACTGATGCT 1690

RESULT 13  
 ADC26254  
 ID ADC26254 standard; DNA; 1502 BP.

AC ADC26254;  
 XX AC  
 DT 18-DEC-2003 (first entry)  
 XX DT  
 DE Human NOV17c DNA.  
 XX DE  
 KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;  
 KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;  
 KW transgenic; human; gene; ds.  
 XX KW  
 OS Homo sapiens.  
 XX OS  
 PN WO2003004687-A2.  
 XX PN  
 PD 16-JAN-2003.  
 XX PD  
 PF 03-JUL-2002; 2002WO-US021361.  
 XX PF  
 PR 05-JUL-2001; 2001US-0303046P.  
 XX PR  
 PR 09-JUL-2001; 2001US-0303828P.  
 XX PR  
 PR 09-JUL-2001; 2001US-0304016P.  
 XX PR  
 PR 11-JUL-2001; 2001US-0304502P.  
 XX PR  
 PR 13-JUL-2001; 2001US-0305262P.  
 XX PR  
 PR 16-JUL-2001; 2001US-0305673P.

60 GTCACTGAGACCCATGGCAAGGAAACAAATAGGAATTCGAAGGAACCTGGGCTAGTTCC 119  
 1 GTCACTGAGACCCATGGCAAGGAAACAAATAGGAATTCGAAGGAACCTGGGCTAGTTCC 60  
 120 CCTCACAGATGACACAGACCCAGCGGGCTTCAGGGCCAGGAGGCACTGCTGAGTG 179  
 61 CCTCACAGATGACACAGACCCAGCGGGCTTCAGGGCCAGGAGGCACTGCTGAGTG 120  
 180 TGACCACTCAGAGAGTGGGTGCGAGGTGGAAGGAGAGAAAGGA- ----CTGGTCTCTGC 234  
 121 TGACCACTCAGAGAGTGGGTGCGAGGTGGAAGGAGAGAAAGGAGCCTTACGAGGAC 180  
 235 TCAGCTCTGCTGAGCTCCCTTCGCGGGCGCTTCGGCTCCTCTTCTTCTTCTTCTTCTTCTTCT 294  
 181 CTCCTCTGACGAGGCTCCTCAACAACATATGTGGCCAGTGTCTATTAAGATCCCAT 240  
 295 CTGTCGGTGTGAATGCCCCACCCCGTACATCAAGGCCCTTT----- 336  
 241 TCACAGGTGGCAAGCTTAGCCCCAGAAAGTCAAGTCACTTGCCTCAGACTCCTACAGCT 300  
 337 ----TACAATGAGTCATGGGAAGGAGCATGGAGCTCCATAGACCCAGACACTCTGA 391  
 301 GAGGGAGCTGGCCCTGGAGGTAAGCTGATATCACTTGGCTCAAGGCCCAAGCTCTAT 360  
 392 CTCGTCTCTGG-----TCTGTGACTGTGTCCATATTCGCC 426  
 361 CTCGTGCTGGTGGCACTAGAGGAGACAAAGAGATTGGCAGAGACTGCTCTGCTGCT 420  
 427 ATCGGTGAGCTTGTGGGAGCTTAATGTG-----AGATGAT 464  
 421 CCTCGTGGCTCCTCCTCGCGGGCGCTTGGGCTCCTCTTCTTCTTCTTCTTCTTCTTCTTCT 480  
 465 TGGAAAGGTTCTTGGGAGGAAAGCACACTTGTGCGCAATATGAGTTTGCATTTCTGC 524  
 481 GGTGGTGAATGCCCCCACCAGGACACTTGTGCGCAATATGAGTTTGCATTTCTGC 540  
 525 TGCAATGCTGATGGCTCTGCTGCTCCAGGACAGGACCTTTGAAATGCTCATTTGGGAGC 584  
 541 TGCAATGCTGATGGCTCTGCTGCTCCAGGACAGGACCTTTGAAATGCTCATTTGGGAGC 600  
 585 CTTTCATCATGGGCAATGATGAGGAGTGGGCTGCGCTTCAAGTGTCTTCCCTACCTCAGTGA 644  
 601 CTTTCATCATGGGCAATGATGAGGAGTGGGCTGCGCTTCAAGTGTCTTCCCTACCTCAGTGA 660  
 645 GATCTCACCCAGGATCCGTGGCTCTCTGCGGAGGAGTGCATGCTTATCTGAT 704  
 661 GATCTCACCCAGGATCCGTGGCTCTCTGCGGAGGAGTGCATGCTTATCTGAT 720  
 705 TGGCGTGTTCACCTGGGACGCTTCTGGGCTGCGCGGAGTGTGGGAAAGGAGAGTACCTG 764  
 721 TGGCGTGTTCACCTGGGACGCTTCTGGGCTGCGCGGAGTGTGGGAAAGGAGAGTACCTG 780  
 765 GCCATACCTGTTGGAGTGAATGTGCTCCTGCGGTGCTCAGCTGCTGAGGCTTCCCTT 824  
 781 GCCATACCTGTTGGAGTGAATGTGCTCCTGCGGTGCTCAGCTGCTGAGGCTTCCCTT 840  
 825 TCTCCCGGACAGCCACGCTACTGCTCTTGGAGAGCACACGAGGCAAGAGCTGTGAA 884  
 841 TCTCCCGGACAGCCACGCTACTGCTCTTGGAGAGCACACGAGGCAAGAGCTGTGAA 900  
 885 AGCCTTCCAAAAGCTTCTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGTCTCGGC 944  
 901 AGCCTTCCAAAAGCTTCTGGGTAAAGCAGACGCTTTCCCAAGAGGTAGAGGAGTCTCGGC 960  
 945 TGAGAGCCAGCTGCAGAGAGGATCCGCTGTGTGCTCGGTGCTGAGGAGTCTTGCAGAGTCC 1004  
 961 TGAGAGCCAGCTGCAGAGAGGATCCGCTGTGTGCTCGGTGCTGAGGAGTCTTGCAGAGTCC 1020  
 1005 CTACGCTCCGCTGCGAGGCTGCTCAGCTGATTTCTCACCATGGCTGCTACCAAGCTCTGAG 1064  
 1021 CTACGCTCCGCTGCGAGGCTGCTCAGCTGATTTCTCACCATGGCTGCTACCAAGCTCTGAG 1080



DB 958 TCATACCAACAGCATCTTTGAAAGCTGGATCCCTCGGCAAGATCCCATAGTCA 1017  
QY 1142 CTTGAGTACAGGGGATCAGAGCTTTGGCTGCGCTCTTCTCTGTTTGGTCAATGAGC 1201  
DB 1018 CTTGAGTACAGGGGATCAGAGCTTTGGCTGCGCTCTTCTCTGTTTGGTCAATGAGC 1077  
QY 1202 ACCTGGGACGAGACCCCTCTCTATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTGGGA 1261  
DB 1078 ACCTGGGACGAGACCCCTCTCTATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTGGGA 1137  
QY 1262 CCCTCACCATCAGCTGACCTCTGAGGACCAACGCGCCCTGGTCCCTACTGAGTATCG 1321  
DB 1138 CCCTCACCATCAGCTGACCTCTGAGGACCAACGCGCCCTGGTCCCTACTGAGTATCG 1197  
QY 1322 TGGGCAATCTGGCCATCATCGCCCTCTTCTGAGTGGGCGCAGGTGGCATCCCTTCACT 1381  
DB 1198 TGGGCAATCTGGCCATCATCGCCCTCTTCTGAGTGGGCGCAGGTGGCATCCCTTCACT 1257  
QY 1382 TGACTGGTGGTCTTCCAGCAATCTCAGGGCGCGCTCTTCTGAGTGGGCGCAGGTGGCATCCCTTCACT 1441  
DB 1258 TGACTGGTGGTCTTCCAGCAATCTCAGGGCGCGCTCTTCTGAGTGGGCGCAGGTGGCATCCCTTCACT 1317  
QY 1442 TCAACTGGCTCTCCAACTTTTCTGTTGGGCTCTCTTCCCAATTCATTCAGAAAAGTCTGG 1501  
DB 1318 TCAACTGGCTCTCCAACTTTTCTGTTGGGCTCTCTTCCCAATTCATTCAGAAAAGTCTGG 1377  
QY 1502 ACACCTACTGTTCTGAGTCTTCTGCTACAAATTTGATACAGAGTGTCTACTCTGTATT 1561  
DB 1378 ACACCTACTGTTCTGAGTCTTCTGCTACAAATTTGATACAGAGTGTCTACTCTGTATT 1437  
QY 1562 TTGTGCTGCTGAGACCAAAACAGAACCTATGAGAAATCAGCAAAATCAGCAAGGATTTTCCAAA 1621  
DB 1438 TTGTGCTGCTGAGACCAAAACAGAACCTATGAGAAATCAGCAAAATCAGCAAGGATTTTCCAAA 1497  
QY 1622 GGAACAAGACATACCCACCAAGAGAGAAATCGACTGCTGCTGATGCT 1674  
DB 1498 GGAACAAGACATACCCACCAAGAGAGAAATCGACTGCTGCTGATGCT 1550

RESULT 11  
ADB48664  
ID ADB48664 standard; cDNA; 2842 BP.

XX AC ADB48664;  
XX DT 04-DEC-2003 (first entry)  
XX DE Novel human cDNA SEQ ID NO 574.  
XX KW ss; cancer; neurodegenerative disease; human.  
XX OS Homo sapiens.

XX PN US2003104529-A1.  
XX PD 05-JUN-2003.  
XX PF 04-JAN-2002; 2002US-00037270.  
XX PR 21-JAN-2000; 2000US-00488725.  
XX PR 25-APR-2000; 2000US-00552317.  
XX PR 19-JUL-2000; 2000US-00620312.

XX PA (ZHOU/) ZHOU P.  
XX PA (TANG/) TANG Y T.  
XX PA (LIUC/) LIU C.  
XX PA (ASUN/) ASUNDI V.  
XX PA (DRMA/) DRMANAC R T.

XX PI Zhou P, Tang YT, Liu C, Asundi V, Drmanac RT;  
XX WPI; 2003-678194/64.

XX PT New polynucleotide, useful for treating diseases e.g., cancer or  
XX neurodegenerative diseases.  
XX PS Claim 1; SEQ ID NO 574; 99pp; English.  
XX CC The invention relates to a polynucleotide comprising a sequence given in  
CC the specification, or its mature protein-coding portion, or its  
CC complement. The polynucleotide is useful for treating diseases e.g.,  
CC cancer or neurodegenerative diseases and many others listed in the  
CC specification. The present sequence represents a novel human cDNA. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification but was obtained in electronic format directly from USPTO  
CC at seqdata.uspto.gov/sequence.html?DocID=20030104529.

XX SQ Sequence 2842 BP; 698 A; 709 C; 693 G; 733 T; 0 U; 9 Other;

Query Match 56.2%; Score 1317.8; DB 8; Length 2842;  
Best Local Similarity 89.9%; Pred. No. 1.4e-295;  
Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 2;  
QY 2 CGACCCACGGCTCGGCTTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGTGGGT 61  
DB 40 CGACCCACGGCTCG- GCCTTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGTGGGT 98  
QY 62 CACTGAGACCCATGGCAGGAAACAAATAGGAATTCAGGAACTGGGCTAGTCCGC 121  
DB 99 CACTGAGACCCATGGCAGGAAACAAATAGGAATTCAGGAACTGGGCTAGTCCGC 158  
QY 122 TCACAGATGACACAGCAGCCAGCGGCTCCAGGGCCAGGAGGACACTGCTGAGTGTG 181  
DB 159 TCACAGATGACACAGCAGCCAGCGCTCCAGGGCCAGGAGGACACTGCTGAGTGTG 218  
QY 182 ACCACCTGAGAGTGGGGTCCAGGTGGAAGAGAGAAAGGACTGGTCTGCTCGCTCC 241  
DB 219 ACCACCTGAGAGTGGGGTCCAGGTGGAAGAGAGAAAGGACTGGTCTGCTCGCTCC 278  
QY 242 TCCTGGGCTCCCTCGGGGGCGCTTGGGCTCTCTCTCTACGGCTACACCTGTCGG 301  
DB 279 TCGTGGGCTCCCTCGGGGGCGCTTGGGCTCTCTCTCTACGGCTACACCTGTCGG 338  
QY 302 TGGTGAATGCCCCACACCCGTACATCAAGGCGCTTTTACAATGAGTATGGAAAGAGGC 361  
DB 339 TGGTGAATGCCCCACACCC- 357  
QY 362 ATGAGCTCCAAATAGACCCAGACACTCTGACTCTGCTCTGCTGTGTGACTGTGTCATAT 421  
DB 358 - 357  
QY 422 TCGCCATCGGTGACTTGTGGGACGTTAATTTGTGAGATGATTTGAAAGGTTCTTGGGA 481  
DB 358 - 357  
QY 482 GGAAGACACTTTTGTGGGCAATAATGGGTTTGCATTTTCTGCTGCAATTTGCTGATGGCCT 541  
DB 358 GGAAGACACTTTTGTGGGCAATAATGGGTTTGCATTTTCTGCTGCAATTTGCTGATGGCCT 417  
QY 542 GCTCGCTCCAGGACGAGCCTTTGAAATGCTCATTTGTGGGACGCTTTCATCATGGGCATAG 601  
DB 418 GCTCGCTCCAGGACGAGCCTTTGAAATGCTCATCGTGGGACGCTTTCATCATGGGCATAG 477  
QY 602 ATGGAGCGTTCGCCCTCAGTGTGCTCCCATGATGATCTCAGTGTGATCTCACCCAGAGAGA 661  
DB 478 ATGGAGCGTTCGCCCTCAGTGTGCTCCCATGATGATCTCAGTGTGATCTCACCCAGAGAGA 537  
QY 662 TCGTGGCTCTTGGGCGAGGTGACATGCGCATCTTTATCTGCAATTTGGCGTGTTCACCTGGGC 721  
DB 538 TCGTGGCTCTTGGGCGAGGTGACATGCGCATCTTTATCTGCAATTTGGCGTGTTCACCTGGGC 597  
QY 722 AGCTTCTGGGCTGCGCGAGCTGCTGGGAAAGAGAGATACCTGGGCAATACCTGTTTGGAG 781  
DB 598 AGCTTCTGGGCTGCGCGAGCTGCTGGGAAAGAGAGATACCTGGGCAATACCTGTTTGGAG 657





SQ Sequence 1710 BP; 393 A; 444 C; 442 G; 431 T; 0 U; 0 Other;  
 Query Match 59.7%; Score 1399; DB 9; Length 1710;  
 Best Local Similarity 93.6%; Pred. No. 1.8e-314;  
 Matches 1511; Conservative 0; Mismatches 5; Indels 99; Gaps 1;  
 60 GTCACAGACCCATGCGAAGAACAAATAGGAATCCAAAGAACTGGGCTAGTTCC 1119  
 1 GTCACAGACCCATGCGAAGAACAAATAGGAATCCAAAGAACTGGGCTAGTTCC 60  
 120 CCTCACAGATGACACACAGCGCGGCTCCAGGCGCAGGAGGAGCACTGCTGGAGTG 179  
 61 CCTCACAGATGACACACAGCGCGGCTCCAGGCGCAGGAGGAGCACTGCTGGAGTG 120  
 180 TGACACCTGAGGAGTGGGGTCCAGGTGGAGGAGAGAAAGCACTGGTCTGCTCGCT 239  
 121 TGACACCTGAGGAGTGGGGTCCAGGTGGAGGAGAGAA----- 162  
 240 CCTGCTGCGCTCCCTGCGGGGCGCTTCGGGCTCCTCTCTACGGCTACAACTGTC 299  
 163 ----- 162  
 300 GGTGAGTAAAGCGCCCGTACATCAAGGCGCTTTTACATGAGTCATGGGAAGAG 359  
 163 -----GTACATCAAGGCGCTTTTACATGAGTCATGGGAAGAG 201  
 360 GCATGAGAGTCCAAATAGACCCAGACATCTGACTCTGCTGTGCTGTGACTGTGTCAT 419  
 202 GCATGAGAGTCCAAATAGACCCAGACATCTGACTCTGCTGTGCTGTGACTGTGTCAT 261  
 420 ATTGCCATCGGTGACTTTGGGAGCGTTAATTGTGAAGATGATGGAAAGTTCCTGG 479  
 262 ATTGCCATCGGTGACTTTGGGAGCATTAATTGTGAAGATGATGGAAAGTTCCTGG 321  
 480 GAGGAGACACTTTGCTGGGCAATATATGGGTTTGGCAATTTCTGCTGATGCTGATGGC 539  
 322 GAGGAGACACTTTGCTGGGCAATATATGGGTTTGGCAATTTCTGCTGATGCTGATGGC 381  
 540 CTGCTCGCTCCAGGAGAGCGCTTTGAATGCTCATTTGCGGAGCGCTTCATCATGGGCAT 599  
 382 CTGCTCGCTCCAGGAGAGCGCTTTGAATGCTCATTTGCGGAGCGCTTCATCATGGGCAT 441  
 600 AGATGAGGCGCTGCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCAAGGA 659  
 442 AGATGAGGCGCTGCGCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCAAGGA 501  
 660 GATCGTGCTCTCTGGGCGAGTGACTGCGCATCTTTATCTGCAATGGCGTGTCTACTGG 719  
 502 GATCGTGCTCTCTGGGCGAGTGACTGCGCATCTTTATCTGCAATGGCGTGTCTACTGG 561  
 720 GCAGCTTCTGGGCTGCGGAGCTGCTGGGAAGGAGAGTACCTGGCCATACCTGTTGG 779  
 562 GCAGCTTCTGGGCTGCGGAGCTGCTGGGAAGGAGAGTACCTGGCCATACCTGTTGG 621  
 780 AGTGATTTGGTCCCTGCGCTGTCCAGTGTGCTGAGCGCTTCCTTTCTCCCGGACAGGCC 839  
 622 AGTGATTTGGTCCCTGCGCTGTCCAGTGTGCTGAGCGCTTCCTTTCTCCCGGACAGGCC 681  
 840 ACGCTACCTGCTCTTGGGAAGCACACAGGAGAGAGTGTGAAGCGCTTCCAAAGCTT 899  
 682 ACGCTACCTGCTCTTGGGAAGCACACAGGAGAGAGTGTGAAGCGCTTCCAAAGCTT 741  
 900 CTTGGGTAAAGCAGAGCTTTCCAAAGAGTGTGAAGAGTCTGCTGAGAGCGCAGTGCA 959  
 742 CTTGGGTAAAGCAGAGCTTTCCAAAGAGTGTGAAGAGTCTGCTGAGAGCGCAGTGCA 801  
 960 GAGGAGCATCCGCTGGTGTCCGCTGAGAGTGTGAGAGTCTGCTGAGAGCGCAGTGCA 1019  
 802 GAGGAGCATCCGCTGGTGTCCGCTGAGAGTGTGAGAGTCTGCTGAGAGCGCAGTGCA 861  
 1020 GGTGGTCACTGATTTGTCACCATGGCTGCTACAGCTCTGTGGCTCAATGCAATTTG 1079  
 862 GGTGGTCACTGATTTGTCACCATGGCTGCTACAGCTCTGTGGCTCAATGCAATTTG 921

QY 1080 GTTCTATACCAACAGCATCTTTTGGAAAGCTGGGATCCTCCGGAAGAGATCCCATACGT 1139  
 DB 922 GTTCTATACCAACAGCATCTTTTGGAAAGCTGGGATCCTCTGGCAAGAGATCCCATACGT 981  
 QY 1140 CACCTTGGAGTACAGGGGGCATCGAGACTTTGGCTGCCGTCTTCTCTGCTTGGTTCATTGA 1199  
 DB 982 CACCTTGGAGTACAGGGGGCATCGAGACTTTGGCTGCCGTCTTCTCTGCTTGGTTCATTGA 1041  
 QY 1200 GACCTGGGACGGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGG 1259  
 DB 1042 GACCTGGGACGGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCTTTGG 1101  
 QY 1260 GACCTCTACCATCAGCTGACCTGACCCCTCAGACACACGCCCTGGGCTCCCTACTCAGATAT 1319  
 DB 1102 GACCTCTACCATCAGCTGACCTGACCCCTCAGACACACGCCCTGGGCTCCCTACTCAGATAT 1161  
 QY 1320 CTTGGGCAATCTGGGCAATCATCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTCAT 1379  
 DB 1162 CTTGGGCAATCTGGGCAATCATCGCTCTTTCTGAGTGGGCGAGGTGGCATCCCGTTTCAT 1221  
 QY 1380 CTTGACTGGTGGTCTTCTCCAGCAATCTCAGCGGCGGCTGCTTCAATTCAGGGCAC 1439  
 DB 1222 CTTGACTGGTGGTCTTCTCCAGCAATCTCAGCGGCGGCTGCTTCAATTCAGGGCAC 1281  
 QY 1440 CTTGACTGGTCTCCAACTTTGCTGGTGGGCTCCTCTTCCCATTCATTCAGAAAGTCT 1499  
 DB 1282 CTTGACTGGTCTCCAACTTTGCTGGTGGGCTCCTCTTCCCATTCATTCAGAAAGTCT 1341  
 QY 1500 GACACCTACTCTTCTCTAGTCTTTGCTACAAATTTGATACAGGTGCTATCTACTGTA 1559  
 DB 1342 GACACCTACTCTTCTCTAGTCTTTGCTACAAATTTGATACAGGTGCTATCTACTGTA 1401  
 QY 1560 TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAGAAATCAGCAGGCAATTTTCCAA 1619  
 DB 1402 TTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAGAAATCAGCAGGCAATTTTCCAA 1461  
 QY 1620 AAGGAAACAAAGCATACCCACCAAGAGAGAAATCGACTCAGTGTGCTGATGCT 1674  
 DB 1462 AAGGAAACAAAGCATACCCACCAAGAGAGAAATCGACTCAGTGTGCTGATGCT 1516  
 RESULT 9  
 AAKS1571  
 ID AAKS1571 standard; cDNA; 1704 BP.  
 XX AC AAKS1571;  
 XX DT 06-NOV-2001 (first entry)  
 XX DE Human polynucleotide SEQ ID NO 116.  
 XX KW Human; cytokines; cell proliferation; cell differentiation; gene therapy;  
 XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;  
 XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;  
 XX KW nervous system disorder; arthritis; inflammation; ss.  
 OS Homo sapiens.  
 PN XX  
 FN XX  
 PD 09-AUG-2001.  
 XX PF 05-FEB-2001; 2001WO-US004098.  
 XX PR 03-FEB-2000; 2000US-00496914.  
 XX PR 27-APR-2000; 2000US-00560875.  
 XX PR 20-JUN-2000; 2000US-00598075.  
 XX PR 19-JUL-2000; 2000US-00620325.  
 XX PR 01-SEP-2000; 2000US-00654936.  
 XX PR 15-SEP-2000; 2000US-00663561.  
 XX PR 20-OCT-2000; 2000US-00693325.  
 XX PR 30-NOV-2000; 2000US-00728422.



QY 744 GCTGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTCAATTGGTCCCTGCCGTTGT 803  
Db 728 GCTGGAAAGAGAGTACCTGGCCATACCTGTTTGGAGTCAATTGGTCCCTGCCGTTGT 787  
QY 804 CCAGCTGCTGAGCCTCCCTTCTCCGGACAGCCACGCTACCTGCTCTTGGAGAGCA 863  
Db 788 CCAGCTGCTGAGCCTCCCTTCTCCGGACAGCCACGCTACCTGCTCTTGGAGAGCA 847  
QY 864 CAACGAGCAGAGCTGTGAAGCCTTCCAAAGCTTCTGGGTAAAGCAGAGCTTCCCA 923  
Db 848 CAACGAGCAGAGCTGTGAAGCCTTCCAAAGCTTCTGGGTAAAGCAGAGCTTCCCA 907  
QY 924 AGAGTAGAGAGTCTGCTGCTGAGAGCCACGCTGAGAGAGCATCCGCTGCTGCTGCT 983  
Db 908 AGAGTAGAGAGTCTGCTGCTGAGAGCCACGCTGAGAGAGCATCCGCTGCTGCTGCT 967  
QY 984 GCTGGAGCTGCTGAGAGCTCCCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043  
Db 968 GCTGGAGCTGCTGAGAGCTCCCTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1027  
QY 1044 GGCCTGTACAGCTGTGGGCTCAATGCAATTTGGTCTTATCAACAGAGCATCTTTGG 1103  
Db 1028 GGCCTGTACAGCTGTGGGCTCAATGCAATTTGGTCTTATCAACAGAGCATCTTTGG 1087  
QY 1104 AAAAGCTGGATCCCTCCGCGCAAGATCCATACGCTGCTGCTGCTGCTGCTGCTGCT 1163  
Db 1088 AAAAGCTGGATCCCTCCGCGCAAGATCCATACGCTGCTGCTGCTGCTGCTGCTGCT 1147  
QY 1164 GACTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1223  
Db 1148 GACTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1207  
QY 1224 CATTGTGGCTTTGGGCTCATNGGCTCTTCTTTTGGGACCTTCACATCAGCTGACCT 1283  
Db 1208 CATTGTGGCTTTGGGCTCATNGGCTCTTCTTTTGGGACCTTCACATCAGCTGACCT 1267  
QY 1284 GCAGGACCGCCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343  
Db 1268 GCAGGACCGCCCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1327  
QY 1344 CTCTTTCTGAGTGGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1403  
Db 1328 CTCTTTCTGAGTGGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1387  
QY 1404 ATCTAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1463  
Db 1388 ATCTAGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1447  
QY 1464 TGTGGGCTCTCTTCCGCTTCAATTCAGAAAGTCTGGACACTGCTGCTGCTGCTGCTGCT 1523  
Db 1448 TGTGGGCTCTCTTCCGCTTCAATTCAGAAAGTCTGGACACTGCTGCTGCTGCTGCTGCT 1507  
QY 1524 TGTACAAATTGTATCACAGGTGCTATCTACCTGCTATTTTGGCTGCTGAGACCAAAA 1583  
Db 1508 TGTACAAATTGTATCACAGGTGCTATCTACCTGCTATTTTGGCTGCTGAGACCAAAA 1567  
QY 1584 CAGAACCTATGAGAAATCAGCCAGGCAATTTCCAAAGGAAACAAAGCATACCCACCA 1643  
Db 1568 CAGAACCTATGAGAAATCAGCCAGGCAATTTCCAAAGGAAACAAAGCATACCCACCA 1627  
QY 1644 AGAGAAATCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1674  
Db 1628 AGAGAAATCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1658

RESULT 8  
ADC26250  
ID ADC26250 standard; DNA; 1710 BP.  
XX AC ADC26250;  
XX AC  
XX AC  
DT 18-DEC-2003 (first entry)

XX DE Human NOV17a DNA.  
XX KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;  
KW KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;  
XX KW transgenic; human; gene; ds.  
XX OS Homo sapiens.  
XX PN WO2003004687-A2.  
XX PD 16-JAN-2003.  
XX PF 03-JUL-2002; 2002WO-US021361.  
XX PR 05-JUL-2001; 2001US-0303046P.  
PR 09-JUL-2001; 2001US-0303828P.  
PR 09-JUL-2001; 2001US-0304016P.  
PR 11-JUL-2001; 2001US-0304502P.  
PR 13-JUL-2001; 2001US-0305262P.  
PR 16-JUL-2001; 2001US-0305673P.  
PR 17-JUL-2001; 2001US-0306085P.  
PR 24-JUL-2001; 2001US-0307536P.  
PR 27-JUL-2001; 2001US-0308288P.  
PR 30-JUL-2001; 2001US-0308877P.  
PR 01-AUG-2001; 2001US-0309255P.  
PR 17-AUG-2001; 2001US-031328P.  
PR 12-SEP-2001; 2001US-0318711P.  
PR 19-SEP-2001; 2001US-0323380P.  
PR 21-SEP-2001; 2001US-0323969P.  
PR 04-JAN-2002; 2002US-0345022P.  
PR 04-JAN-2002; 2002US-0345038P.  
PR 28-FEB-2002; 2002US-0361172P.  
PR 01-MAR-2002; 2002US-0360814P.  
PR 01-MAR-2002; 2002US-0360830P.  
PR 01-MAR-2002; 2002US-0361133P.  
PR 01-MAR-2002; 2002US-0361147P.  
PR 05-MAR-2002; 2002US-0361677P.  
PR 02-APR-2002; 2002US-0363637P.  
PR 12-APR-2002; 2002US-0372326P.  
PR 16-APR-2002; 2002US-0372990P.  
PR 19-APR-2002; 2002US-0373881P.  
PR 19-APR-2002; 2002US-0373921P.  
PR 02-JUL-2002; 2002US-00188186.  
XX PA (CURA-) CURAGEN CORP.  
XX PI Anderson DM, Berghs C, Boldog FL, Burgess CE, Caeman SJ;  
PI Catterton E, Edinger S, Eissen AJ, Ellerman K, Gerlach V, Gorman L;  
PI Guo X, Jeffers M, Kikuda R, Li L, Malyankar UM, Miller CE;  
PI Padigaru M, Patturajan M, Pena CE, Rastelli L, Shenoy S;  
PI Shimkets RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ;  
PI Vernet CAM, Voss EZ, Zhong M;  
XX WPI; 2003-221607/21.  
XX P-PSDB; ADC26251.  
XX PT New isolated NOVX polypeptide, useful for determining the presence of, or  
XX PT predisposition to a disease associated with altered levels of expression  
XX PT of the polypeptide, and for treating or preventing cancer.  
XX PS Claim 20; SEQ ID NO 75; 478pp; English.  
XX CC The invention relates to a novel isolated NOV polypeptide. The  
XX CC polypeptide of the invention demonstrates cytostatic activity and may be  
XX CC used for determining the presence of, or predisposition to a disease  
XX CC associated with altered levels of expression of the polypeptide,  
XX CC including metabolic disorders, immune disorders, neurodegenerative  
XX CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases  
XX CC and cancer. The polypeptide may also be utilised during gene therapy  
XX CC procedures, vaccine development and transgenic animal production. The  
XX CC current sequence is that of the human NOV DNA of the invention.



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

1

1







QY 1981 AGAGAGAGAGAGAGAGATCTATTCTAAAGCTTAGGCTGTGACGGTGGCTC 2040  
 Db 1981 AGAGAGAGAGAGAGAGATCTATTCTAAAGCTTAGGCTGTGACGGTGGCTC 2040  
 QY 2041 AGCCCTGTAATCCAGCACTTTGGAGGCGGAGGTGGTGGATCGTGAAGTCAAGATT 2100  
 Db 2041 AGCCCTGTAATCCAGCACTTTGGAGGCGGAGGTGGTGGATCGTGAAGTCAAGATT 2100  
 QY 2101 GAGACCATCTCGCTAAACATGTAAGTCTCCTCTCTACTAAATAACAAAATTAGCT 2160  
 Db 2101 GAGACCATCTCGCTAAACATGTAAGTCTCCTCTCTACTAAATAACAAAATTAGCT 2160  
 QY 2161 GAGCATGTGGGGGGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGCGAGGAAATGGC 2220  
 Db 2161 GAGCATGTGGGGGGCGCTGTAGTCCAGCTACTTTGGAGGCTGAGCGAGGAAATGGC 2220  
 QY 2221 GTGAACCCAGGAGGCGGAGCTTGCATGAGCGGAGATCACACCACTCCAGCTGG 2280  
 Db 2221 GTGAACCCAGGAGGCGGAGCTTGCATGAGCGGAGATCACACCACTCCAGCTGG 2280  
 QY 2281 GTGACAGCCAGCACTCCGCTCTCAAAAAAAGGAGGCTGAGCGAGGAAATGGC 2340  
 Db 2281 GTGACAGCCAGCACTCCGCTCTCAAAAAAAGGAGGCTGAGCGAGGAAATGGC 2340  
 QY 2341 CGC 2343  
 Db 2341 CGC 2343

RESULT 4  
 AAD27278  
 ID AAD27278 standard; cDNA; 2032 BP.  
 XX  
 AC AAD27278;  
 XX  
 DT 09-APR-2002 (first entry)  
 XX  
 DE Human transporter and ion channel-25 (TRICH-25) cDNA.  
 XX  
 KW Human; transporter and ion channel-25; TRICH-25; neuroprotective; asthma;  
 KW nontropic; cytostatic; cardiovascular; immunosuppressive; cardiomyopathy;  
 KW antiinflammatory; protein therapy; akinesia; cystic fibrosis; leukaemia;  
 KW Bell's palsy; amyotrophic lateral sclerosis; Alzheimer's disease; cancer;  
 KW amnesia; dementia; myocarditis; Duchenne's muscular dystrophy; AIDS;  
 KW Acquired Immune Deficiency Syndrome; Addison's disease; allergy; angina;  
 KW cell proliferative disorder; psoriasis; cardiac disease; hypertension;  
 KW bradyarrhythmia; gene expression; drug screening; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key  
 FH CDS  
 FT Location/Qualifiers  
 FT 57..1670  
 FT /tag= a  
 FT /product= "Human TRICH-25 protein"  
 FT sig\_peptide 57..254  
 FT /tag= b  
 FT mat\_peptide 255..1667  
 FT /tag= c  
 FT /product= "Mature human TRICH-25 protein"  
 XX  
 FN WO200192304-A2.  
 XX  
 PD 06-DEC-2001.  
 XX  
 PF 25-MAY-2001; 2001WO-US017055.  
 XX  
 PR 26-MAY-2000; 2000US-0208424P.  
 PR 01-JUN-2000; 2000US-0209001P.  
 PR 08-JUN-2000; 2000US-0210589P.  
 PR 16-JUN-2000; 2000US-0212335P.  
 PR 22-JUN-2000; 2000US-0213747P.  
 PR 29-JUN-2000; 2000US-0215391P.

XX (INCV-) INCYTE GENOMICS INC.  
 XX Thornton M, Wallia NK, Yue H, Nguyen DB, Lai P, Gandhi AR;  
 PI Tribouley CM, Yao MG, Ramkumar J, Au-Young J, Lu Y, Tang YT;  
 PI Azimzai Y, Bruns CM, Griffin JA, Yang J, Sanjanwala MS, Raumann BE;  
 PI Lee EA, Hafalia A, Baughn MR, Green BD, Khan FA, Kearney L;  
 PI Elliot VS, Seilhamer JJ, Policky JL, Borowsky ML, Burford N, Ding L;  
 PI Lu DAM, Hillman JL;  
 XX WPI: 2002-122055/16.  
 DR P-PSDB; AAE16788.  
 XX  
 XX New human transporters and ion channels (TRICH) polypeptides useful for  
 PT diagnosing, treating or preventing disorders associated with aberrant  
 PT expression of TRICH.  
 XX  
 PS Claim 11; Page 207-208; 210pp; English.  
 XX  
 CC The invention relates to human transporters and ion channels (TRICH)  
 CC polypeptides and their cDNA molecules. The nucleic acid and polypeptide  
 CC sequences are useful in the diagnosis, treatment, and prevention of  
 CC disorders associated with transport (akinesia, cystic fibrosis, Bell's  
 CC palsy, amyotrophic lateral sclerosis); neurological (Alzheimer's disease,  
 CC amnesia, dementia); muscle (cardiomyopathy, myocarditis, Duchenne's  
 CC muscular dystrophy); immunological (AIDS, Addison's disease, allergies,  
 CC asthma); cell proliferative disorders (cancers, leukaemia, psoriasis);  
 CC cardiac disease (angina, hypertension, or bradyarrhythmia) and in the  
 CC assessment of the effects of exogenous compounds on the expression of  
 CC nucleic acid and amino acid sequences of transporters and ion channels.  
 CC The polynucleotides may be used to detect and quantify gene expression in  
 CC biopsied tissues in which TRICH expression may be correlated with a  
 CC disease, to generate hybridization probes for mapping naturally occurring  
 CC genomic sequence, and in drug screening. The present sequence is human  
 CC TRICH-25 cDNA  
 XX  
 SQ Sequence 2032 BP; 438 A; 557 C; 536 G; 501 T; 0 U; 0 Other;  
 Query Match 70.5%; Score 1651; DB 6; Length 2032;  
 Best Local Similarity 99.7%; Pred. No. 0;  
 Matches 1654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 17 GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG 76  
 Db 1 GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG 60  
 QY 77 CAAGGAACAAATAGGAATTCGAAGAACTGGGCTAGTCCCTCCACAGATGACACCA 136  
 Db 61 CAAGGAACAAATAGGAATTCGAAGAACTGGGCTAGTCCCTCCACAGATGACACCA 120  
 QY 137 GCCACGGCGGCTCCACAGGCGCCAGGGGCACTGTGGAGTGTGACCACTGAGGAGTG 196  
 Db 121 GCCACGGCGGCTCCACAGGCGCCAGGGGCACTGTGGAGTGTGACCACTGAGGAGTG 180  
 QY 197 GGGTGGCAGGTGGAAGGAGAAAGAAAGGACTGTGCTCTGCTGCTGCTGCTGCTGCTGCTG 256  
 Db 181 GGGTGGCAGGTGGAAGGAGAAAGAAAGGACTGTGCTCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 257 CGGGCGCTTCGGCTCCT 316  
 Db 241 CGGGCGCTTCGGCTCCT 300  
 QY 317 CCCGTATACAAAGGCGCTTTTACAATGAGTCACTGGGAAAGAGGCAATGGACGTCCAAATAG 376  
 Db 301 CCCGTATACAAAGGCGCTTTTACAATGAGTCACTGGGAAAGAGGCAATGGACGTCCAAATAG 360  
 QY 377 ACCCAGACACTGACT 436  
 Db 361 ACCCAGACACTGACT 420  
 QY 437 TTGTGGGGAGCTTAATTTGTAAGATGATTTGGAAGGTCTTTGGAGGAGGAGCACTTTTCG 496  
 Db 421 TTGTGGGGAGCTTAATTTGTAAGATGATTTGGAAGGTCTTTGGAGGAGGAGCACTTTTCG 480

PT absence of compound.  
XX Claim 4; Fig 1; 48pp; English.  
PS The invention relates to identifying compounds that modulate the activity  
XX of glucose transporter (GLUTX). The compound identified is useful in the  
CC diagnosis and treatment of disorders associated with GLUTX activity and  
CC aberrant hexose transport. The present sequence is human GLUTX cDNA  
XX  
SQ Sequence 2343 BP; 535 A; 637 G; 621 G; 550 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2343; DB 6; Length 2343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCGACCCACGCTCGCGCCCTTCGCGAGCTCGGGTCCCTGACATGAGCCATCAGCTGGG 60  
DB 1 TCGACCCACGCTCGCGCCCTTCGCGAGCTCGGGTCCCTGACATGAGCCATCAGCTGGG 60  
QY 61 TCACTGAGACCCATGCAAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTCCC 120  
DB 61 TCACTGAGACCCATGCAAGGAAACAAATAGGAATTCAGAGAACTGGGCTAGTCCC 120  
QY 121 CTCACAGATGACACACGACCCGCGCTCCAGGCCCAGGAGGACACTGCTGGAGTGT 180  
DB 121 CTCACAGATGACACACGACCCGCGCTCCAGGCCCAGGAGGACACTGCTGGAGTGT 180  
QY 181 GACCACTGAGAGTGGGTGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
DB 181 GACCACTGAGAGTGGGTGCGAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240  
QY 241 CTGCTGGCTCCCTCGCGGCGCTTCGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 300  
DB 241 CTGCTGGCTCCCTCGCGGCGCTTCGCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT 300  
QY 301 GTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTCAATGAGTCATGGGAAAGAGG 360  
DB 301 GTGGTGAATGCCGCCACCCGCTACATCAAGGCTTTTCAATGAGTCATGGGAAAGAGG 360  
QY 361 CATGAGCGTCCATAGACACCCAGACACTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 420  
DB 361 CATGAGCGTCCATAGACACCCAGACACTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 420  
QY 421 TTCGCCATCGGTGGACTTGTGGGAGCTTAAATGTGAAGATGATTTGGAAGTTCCTGGG 480  
DB 421 TTCGCCATCGGTGGACTTGTGGGAGCTTAAATGTGAAGATGATTTGGAAGTTCCTGGG 480  
QY 481 AGGAACACACTTGTGCGCCATATAGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTG 540  
DB 481 AGGAACACACTTGTGCGCCATATAGGTTTGCATTTCTGCTGCTGCTGCTGCTGCTGCTG 540  
QY 541 TGCTCGCTCCAGCGAGCGCTTTGAAATGCTCATTTGCGGAGCTTTCATCATGGGCATA 600  
DB 541 TGCTCGCTCCAGCGAGCGCTTTGAAATGCTCATTTGCGGAGCTTTCATCATGGGCATA 600  
QY 601 GATGAGGGCTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG 660  
DB 601 GATGAGGGCTCGCCCTCAGTGTGCTCCCATGTACCTCAGTGAGATCTCACCCAGGAG 660  
QY 661 ATCCGTGGCTCTCTGGGCGAGTGAATGCAATCTTATCTGATTTGCGGTGTTCTGCTGG 720  
DB 661 ATCCGTGGCTCTCTGGGCGAGTGAATGCAATCTTATCTGATTTGCGGTGTTCTGCTGG 720  
QY 721 CAGCTTCTGGGCTCGCGAGCTGTGGGAAGGAGAGTACCTGGCCATACCTGTTTGA 780  
DB 721 CAGCTTCTGGGCTCGCGAGCTGTGGGAAGGAGAGTACCTGGCCATACCTGTTTGA 780  
QY 781 GTGATTTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 781 GTGATTTGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
QY 841 CGTACTGCTCTGTGAGAGAGCAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900  
DB 841 CGTACTGCTCTGTGAGAGAGCAACAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

841 CGCTACTGCTCTTGGAGAGCAACAGGAGAGAGCTGTGAAGCCTTCCAAACGTTTC 900  
901 TTGGTAAACGACAGCTTTCCCAAGAGGTAGAGAGTCTGCTGAGAGCCACGTCAG 960  
901 TTGGTAAACGACAGCTTTCCCAAGAGGTAGAGAGTCTGCTGAGAGCCACGTCAG 960  
961 AGAGAGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
961 AGAGAGATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020  
1021 GTGGTCAACGCTGATTTGTACCATGGCTGTACAGCTCTGTGGCTCTCAATGCAATTTGG 1080  
1021 GTGGTCAACGCTGATTTGTACCATGGCTGTACAGCTCTGTGGCTCTCAATGCAATTTGG 1080  
1081 TTCTATACCAACAGCATCTTTTGAAGAGTGGATCCCTCCGCAAGATCCCATACGTC 1140  
1081 TTCTATACCAACAGCATCTTTTGAAGAGTGGATCCCTCCGCAAGATCCCATACGTC 1140  
1141 ACCTTGAGTACAGGGGACATCGAGACTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
1141 ACCTTGAGTACAGGGGACATCGAGACTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1200  
1201 CACCTGGAGCGGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG 1260  
1201 CACCTGGAGCGGAGACCCCTCCTCATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG 1260  
1261 ACCCTCACCATCAGCTGACCTTCAGAGACACAGCCCTCGGGTCCCTACCTGAGTATC 1320  
1261 ACCCTCACCATCAGCTGACCTTCAGAGACACAGCCCTCGGGTCCCTACCTGAGTATC 1320  
1321 GTGGGCACTTCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGTGGCATCCCGTTTCA 1380  
1321 GTGGGCACTTCTGGCCATCATCGCTCTTTCTGAGTGGGCGAGTGGCATCCCGTTTCA 1380  
1381 TTGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTCTCATTTGCAAGAGTCTG 1440  
1381 TTGACTGGTGGTCTTCCAGCAATCTCAGCGGCGGCTGCTCTCATTTGCAAGAGTCTG 1440  
1441 GTCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCCATTTCAATTCAGAAAGTCTG 1500  
1441 GTCAACTGGCTCTCCAACTTTGCTGTTGGGCTCCTCTTCCCATTTCAATTCAGAAAGTCTG 1500  
1501 GACACCTACTGTTTCTTAGTCTTTGTACAAATTTGTATCACAGTGTCTATCTACCTGTAT 1560  
1501 GACACCTACTGTTTCTTAGTCTTTGTACAAATTTGTATCACAGTGTCTATCTACCTGTAT 1560  
1561 TTGCTGTGCTGAGACCAAAACAGAACTATGAGAAATCAGCCAGGCAATTTTCCAAA 1620  
1561 TTGCTGTGCTGAGACCAAAACAGAACTATGAGAAATCAGCCAGGCAATTTTCCAAA 1620  
1621 AGGAACAAAGCATACCCACAGAGAGAAATGCACTCAGGTGTCTGATGCTCTGCTGCT 1680  
1621 AGGAACAAAGCATACCCACAGAGAGAAATGCACTCAGGTGTCTGATGCTCTGCTGCT 1680  
1681 TCCTTCTCTTCTACTACTCCGAAATACAGCTGGATTCAGCTGCGCCACACCCAGCC 1740  
1681 TCCTTCTCTTCTACTACTCCGAAATACAGCTGGATTCAGCTGCGCCACACCCAGCC 1740  
1741 ACCAAAAAGAACACCCATTTGTAACGCTCATGTGGTATTTCTCTCAACCTGGATGACCT 1800  
1741 ACCAAAAAGAACACCCATTTGTAACGCTCATGTGGTATTTCTCTCAACCTGGATGACCT 1800  
1801 TCCCTTATCTTCTCTCTGAGAACCAAGTCAATGATGTGAGAACAGAGCTTTGATTT 1860  
1801 TCCCTTATCTTCTCTCTGAGAACCAAGTCAATGATGTGAGAACAGAGCTTTGATTT 1860  
1861 TGGAGACATGGGTTTGAATTCAGTCAATCTTTTATTTTCAAGAAATTTTAAACAGT 1920  
1861 TGGAGACATGGGTTTGAATTCAGTCAATCTTTTATTTTCAAGAAATTTTAAACAGT 1920  
1921 ACTGACATGCTCCCATATGTTGTTTACCCCTGCTTATACAAATGGGAGGAGAGAGAG 1980  
1921 ACTGACATGCTCCCATATGTTGTTTACCCCTGCTTATACAAATGGGAGGAGAGAGAG 1980



QY	1981	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTCAAGTCTAGGCTGTGCA CGTGGCTC	2040
Db	1981	AGAGAGAGAGAGAGATGCTATTCTTAAAGCTTCAAGTCTAGGCTGTGCA CGTGGCTC	2040
QY	2041	AGCCTGTAAATCCCGACATTTGGGAGGCCGAGGTGGTGGATCGTAGGTCAGGAGATT	2100
Db	2041	AGCCTGTAAATCCCGACATTTGGGAGGCCGAGGTGGTGGATCGTAGGTCAGGAGATT	2100
QY	2101	GAGACCATCTGGCTAAACATGTGAAAATCTCCCTCTCTACTATAAATAACAAAAATTAGCT	2160
Db	2101	GAGACCATCTGGCTAAACATGTGAAAATCTCCCTCTCTACTATAAATAACAAAAATTAGCT	2160
QY	2161	GAGCATGGTGGGGGGCGCTGTAGTCCAGCTACTTGGGAGGCTGAGGAGGAGATGGC	2220
Db	2161	GAGCATGGTGGGGGGCGCTGTAGTCCAGCTACTTGGGAGGCTGAGGAGGAGATGGC	2220
QY	2221	GTGAACCCAGGAGGCGGAGCTTTGCAGTGTAGCCAGCTACTTGGGAGGCTGAGGAGGAGATGGC	2280
Db	2221	GTGAACCCAGGAGGCGGAGCTTTGCAGTGTAGCCAGCTACTTGGGAGGCTGAGGAGGAGATGGC	2280
QY	2281	GTGACAGACCGAGCTCCGTCTCAAAAAAATAAAAAAATAAAAAAAGGCGGC	2340
Db	2281	GTGACAGACCGAGCTCCGTCTCAAAAAAATAAAAAAATAAAAAAAGGCGGC	2340
QY	2341	CGC 2343	
Db	2341	CGC 2343	
RESULT 3			
AAD27918			
ID	AAD27918 standard; cDNA; 2343 BP.		
XX	AAD27918;		
XX	18-JUN-2002 (first entry)		
DE	Human glucose transporter protein GLUTX cDNA.		
XX	Glucose transporter; GLUTX; hexose; human; ss.		
XX	Homo sapiens.		
XX	Location/Qualifiers		
FT	73..1764		
FT	/*tag= a		
FT	/product= "Glucose transporter, GLUTX"		
FT	2337..2339		
FT	/*tag= b		
FT	/note= "Nucleotides GGC are present at this location in the sequence shown as SEQ ID NO:1 in Sequence Listing of the specification"		
XX			
PN	US6346374-B1.		
XX			
PD	12-FEB-2002.		
XX			
PF	05-JUL-2000; 2000US-00610417.		
XX			
PR	26-FEB-1998; 98US-00031392.		
PR	26-APR-1999; 99US-00299549.		
XX			
PA	(MILL-) MILLENNIUM PHARM INC.		
XX			
PI	Tartaglia LA, Weng X;		
XX			
DR	WPI; 2002-224989/28.		
DR	P-PSDB; XAE14611.		
XX			
PT	Identifying compound modulating activity of glucose transporter, useful for treating disorders mediated by transporter, comprises contacting polypeptide with test compound and comparing activity in presence and		



QY 1981 AGAGAGAGAGAGAGATCTATTCTTAAAGCTTGAAGTCTAGGCTGTGCACGGTGGCTC 2040  
DB 1981 AGAGAGAGAGAGAGATCTATTCTTAAAGCTTGAAGTCTAGGCTGTGCACGGTGGCTC 2040  
QY 2041 ACSCCTGTATCCACACTTTGGGAGGCGGAGGTGGGATCGTAGGTGAGGATTT 2100  
DB 2041 ACSCCTGTATCCACACTTTGGGAGGCGGAGGTGGGATCGTAGGTGAGGATTT 2100  
QY 2101 GAGACCATCTGGCTAACATGTTGAACTCCCTCTCTACTATAAATAACAAAAATTAGCT 2160  
DB 2101 GAGACCATCTGGCTAACATGTTGAACTCCCTCTCTACTATAAATAACAAAAATTAGCT 2160  
QY 2161 GAGCATGTGGGGGGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGAGAGATGGC 2220  
DB 2161 GAGCATGTGGGGGGCGCTGTAGTCCAGCTACTTGGAGGCTGAGGAGAGATGGC 2220  
QY 2221 GTGAACCCAGGAGGCGGAGCTTGCAGTGAGCGGAGATCACACCACACTCCAGCGTGG 2280  
DB 2221 GTGAACCCAGGAGGCGGAGCTTGCAGTGAGCGGAGATCACACCACACTCCAGCGTGG 2280  
QY 2281 GTGACAGAGCCAGCTCCGTCTCAAAAAAATAAAAAAATAAAAAAAGGCGGCGC 2340  
DB 2281 GTGACAGAGCCAGCTCCGTCTCAAAAAAATAAAAAAATAAAAAAAGGCGGCGC 2340  
QY 2341 CGC 2343  
DB 2341 CGC 2343

RESULT 2  
AAC62143  
ID AAC62143 standard; cDNA; 2343 BP.  
XX AAC62143;  
AC AAC62143;  
XX  
XX 06-MAR-2001 (first entry)  
XX cDNA encoding a human GLUTX polypeptide.  
XX Human; GLUTX; glucose transporter; hexose transport;  
KW tissue-specific marker; ss.  
XX Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH 73..1764  
FT /\*tag= a  
FT /product= "GLUTX"  
XX  
XX US6136547-A.  
XX  
XX 24-OCT-2000.  
XX  
XX 26-APR-1999; 99US-00299549.  
XX  
XX 26-FEB-1998; 98US-00031392.  
XX  
XX (MILL-) MILLENNIUM PHARM INC.  
XX  
XX Weng X, Tartaglia LA;  
XX WPI; 2001-040236/05.  
XX P-PSDB; AAB30521.  
XX  
XX New human glucose transporter protein and nucleic acid encoding the  
PT protein, useful in the diagnosis and treatment of disorders associated  
PT with aberrant hexose transport.  
XX  
XX Example; Fig 1A-E; 48pp; English.  
XX  
XX The present sequence encodes a human GLUTX polypeptide. GLUTX is a  
CC glucose transporter protein. The GLUTX polypeptides and polynucleotides  
CC are useful in the diagnosis and treatment of disorders associated with

CC aberrant hexose transport. The GLUTX polypeptide is useful for  
CC determining whether a given disorder is associated with aberrant  
CC expression of GLUTX or activity of GLUTX. The nucleic acid molecule and  
CC the GLUTX polypeptide are useful as diagnostic or therapeutic agents, or  
CC they can be used to generate antibodies or identify small molecules that,  
CC in turn are clinically useful. The GLUTX nucleic acid molecules are  
CC useful for identifying the chromosomal location of GLUTX and as tissue-  
CC specific markers  
XX  
SQ Sequence 2343 BP; 535 A; 637 C; 621 G; 550 T; 0 U; 0 Other;  
Query Match 100.0%; Score 2343; DB 5; Length 2343;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 2343; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 TCAGCCACACGCTCCGGCTTGGCAGAGTCTGGGGTCCCTTGGACTGAGCCATCAGCTGGG 60  
DB 1 TCAGCCACACGCTCCGGCTTGGCAGAGTCTGGGGTCCCTTGGACTGAGCCATCAGCTGGG 60  
QY 61 TCATGTAGACCCATGGCAAGAAACAAATAGGAATTCAGAGAACTCGGGCTAGTTCCC 120  
DB 61 TCATGTAGACCCATGGCAAGAAACAAATAGGAATTCAGAGAACTCGGGCTAGTTCCC 120  
QY 121 CTACAGATGACACACGCGGCTCCAGGGCCAGGGAGGCACTGCTGGAGTGT 180  
DB 121 CTACAGATGACACACGCGGCTCCAGGGCCAGGGAGGCACTGCTGGAGTGT 180  
QY 181 GACCACCTGAGGAGTGGGGTGCAGAGTGAAGAGAAAGAGCACTGCTCGCTC 240  
DB 181 GACCACCTGAGGAGTGGGGTGCAGAGTGAAGAGAAAGAGCACTGCTCGCTC 240  
QY 241 CTGCTGGCTCCCTCGCGGGCGCTTCGGGTCTCTTCCTTACGCTCAACCTGTG 300  
DB 241 CTGCTGGCTCCCTCGCGGGCGCTTCGGGTCTCTTCCTTACGCTCAACCTGTG 300  
QY 301 GTGTGTAATGCCCCACCCCTGATCATCAAGCCCTTTTACATGAGTATGGAAGAAGG 360  
DB 301 GTGTGTAATGCCCCACCCCTGATCATCAAGCCCTTTTACATGAGTATGGAAGAAGG 360  
QY 361 CATGAGCTCCATAGACCCAGACACTCTGCTCTGCTCTGCTCTGCTCTGCTCATA 420  
DB 361 CATGAGCTCCATAGACCCAGACACTCTGCTCTGCTCTGCTCTGCTCTGCTCATA 420  
QY 421 TTGCGCATCGGTGAGCTTTGGGGAGCGTTAAATGGAAGATGATTGGAAGATTTCTGGG 480  
DB 421 TTGCGCATCGGTGAGCTTTGGGGAGCGTTAAATGGAAGATGATTGGAAGATTTCTGGG 480  
QY 481 AGGAAGCACACTTTGCTGGCCAAATAATGGGTTTGCATTTCTGCTGCTGCTGATGGCC 540  
DB 481 AGGAAGCACACTTTGCTGGCCAAATAATGGGTTTGCATTTCTGCTGCTGCTGATGGCC 540  
QY 541 TGCTCGCTCCAGGAGAGCGCTTTGAAATGCTCATTTGTGGAGCGCTTCATCATGGGCATA 600  
DB 541 TGCTCGCTCCAGGAGAGCGCTTTGAAATGCTCATTTGTGGAGCGCTTCATCATGGGCATA 600  
QY 601 GATGAGCGTGGCGCTCAGTGTCTCCCATGATGATGATGATGATGATGATGATGATGATG 660  
DB 601 GATGAGCGTGGCGCTCAGTGTCTCCCATGATGATGATGATGATGATGATGATGATGATG 660  
QY 661 ATCCGTGGCTCTGCGGGCAGTGTGCTGCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTG 720  
DB 661 ATCCGTGGCTCTGCGGGCAGTGTGCTGCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTG 720  
QY 721 CAGCTTCTGGGCTGCGCGAGCTGTGGAAAGAGAGATGATGATGATGATGATGATGATGATG 780  
DB 721 CAGCTTCTGGGCTGCGCGAGCTGTGGAAAGAGAGATGATGATGATGATGATGATGATGATG 780  
QY 781 GTGATGTGTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
DB 781 GTGATGTGTGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840  
QY 841 CGCTACCTGCTTGGAGAGCACAACGAGGAGAGAGCTGTGAAAGCTTCGAAACGCTTC 900



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 21:02:47 ; Search time 922 Seconds  
(without alignments)  
10795.584 Million cell updates/sec

Title: US-09-981-947B-1

Perfect score: 2343

Sequence: 1 tcgaccacgcgtccggcct.....aaaaaaaaagggcgccgc 2343

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2343	100.0	2343	2	Aaz11502 Human glu
2	2343	100.0	2343	5	Aac62143 cDNA enco
3	2343	100.0	2343	6	Aad27918 Human glu
4	1851	70.5	2032	6	Aad27278 Human tra
5	1644.6	70.2	1851	4	Abn09154 Human GLU
6	1644.6	70.2	1851	4	Aak52555 Human pol
7	1643	70.1	1851	4	Aai160472 Human pol
8	1399	59.7	1710	9	Adc26250 Human NOV
9	1317.8	56.2	1704	4	Aak51571 Human pol
10	1317.8	56.2	2842	4	Aai58686 Novel hum
11	1317.8	56.2	2842	8	Adb48664 Novel hum
12	1232.8	52.6	1757	9	Adc26252 Human NOV
13	1229	52.5	1502	9	Adc26254 Human NOV
14	1187.6	50.7	1267	9	Adc26264 Human NOV
15	1141	48.7	2216	4	Aad09557 Human tra
16	1133.8	48.4	1189	9	Adc26260 Human NOV
17	1130.6	48.3	1189	9	Adc26258 Human NOV
18	1130.6	48.3	1189	9	Adc26256 Human NOV
19	915	39.1	1087	9	Adc26256 Human NOV
20	430.8	18.4	472	8	Ach43430 Human foe
21	427.6	18.3	1536	9	Adn01366 Human TCH
22	427.6	18.3	1539	9	Acc42990 Human glu
23	427.6	18.3	1563	9	Adn01388 Human var

24	427.6	18.3	1552	9	ADD01378 Human TCH
25	424.4	18.1	1539	6	Aal44673 Human tra
26	421.6	18.0	2169	7	Abt42242 Toxicity
27	421.6	18.0	2169	9	Adb53334 Primary r
28	416.4	17.8	2218	6	Abi164747 Lung canc
29	416.4	17.8	2218	6	Abi165386 Lung canc
30	416.4	17.8	2218	6	Abv78029 Hypoxia-r
31	416.4	17.8	2218	7	Acc72731 Human can
32	410	17.5	7434	6	Aad28386 Human che
33	407.8	17.4	7434	6	Aad28387 Human che
34	376	16.0	1607	6	Abss2105 Human glu
35	354	15.1	1461	7	Acc05717 cDNA enco
36	354	15.1	2064	7	Aad53978 Human cdn
37	348	14.9	2222	5	Aaf93832 Human sbg
38	335.2	14.3	1491	6	Abi157737 Human sbg
39	334.8	14.3	1501	6	Aal44681 Human tra
40	333.2	14.2	2135	6	Abt31931 Human bre
41	330.2	14.1	2778	7	Abx74453 Human cdn
42	330.2	14.1	2854	7	Ada28225 Human lun
43	330.2	14.1	2856	8	Ada28225 Human lun
44	330.2	14.1	2893	5	Aaf98714 Human lac
45	330.2	14.1	3366	6	Abz35362 Human gen

#### ALIGNMENTS

RESULT 1

Aaz11502

ID AAZ11502 standard; cDNA; 2343 BP.

XX AC AAZ11502;

DT 05-NOV-1999 (first entry)

XX Human glucose transporter protein GLUTX encoding cDNA.

XX Human; glucose transporter; GLUTX; hexose transport; gene therapy;

XX Human; glucose mapping; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

FT CDS 73..1764

FT /\*tag= a

FT /product= "GLUTX"

XX US942398-A.

XX 24-AUG-1999.

XX 26-FEB-1998; 98US-00031392.

XX 26-FEB-1998; 98US-00031392.

XX (WILL-) MILLENNIUM PHARM INC.

XX Tartaglia LA, Weng X;

XX WPI; 1999-526192/44.

XX P-PSDB; AAY27286.

XX A novel mammalian transporter protein and the gene encoding it, are

XX useful in the diagnosis and treatment of disorders associated with

XX aberrant sugar transport.

XX Claim 2; Fig 1A-E; 48pp; English.

XX This cDNA encodes a human glucose transporter (GLUTX) protein. The

XX protein can be expressed by standard recombinant methodology. GLUTX

XX nucleic acids are useful as hybridization probes for detecting the

XX presence of GLUTX DNA in a sample, useful for diagnosing conditions

XX associated with aberrant expression levels of GLUTX. The GLUTX gene is

Qy	956	TGCAGAGGAGCATCCGCTCGGTGTCGGTGTGGAGCTGTGAGAGCTCCCTAGTCCGCT	1015
Db	785	CCGAGCGCGGAGGGCCACCTGTCTGTGTGCACCTCTGTGCCCTCGGTCCTCGCT	844
Qy	1016	GGCAGGTGGTCACCGTGAATGTACCATGGCTGTACAGCTCTGTGGCTCAATGCAA	1075
Db	845	GGCAGCTCCTCTCCATCATCGTGTCTATGGCCGCCAGCAGCTGTGGGGCATCAATGCCA	904
Qy	1076	TTTGGTTCTATACCAACAGCATCTTTGAAAAAGTGGATCCCTCCGGGCAAGATCCCAT	1135
Db	905	TCAACTACTATGCGGACACCATCTACATCTGGGGCGTGGAGCGGCTCACTCCCAAT	964
Qy	1136	ACGTCACTTTGAGTACAGGGGCGCATCGAGACTTTGGCTGCCGTCTTCTCTGGTTGGTCA	1195
Db	965	ATGTAACGGTGGGCTCTGGGCTCGTCAACATAGTAGATGACATCACCTCGGTGTCTTG	1024
Qy	1196	TTGAGCACTGGGACGAGACCCCTCTCTCATTTGGTGGCTTTGGGCTCATGGGCTCTTCT	1255
Db	1025	TGGAGCGGCTGGGACGGCGCACCTCTCTGTGGCCGGCTACGGCATCTGGGGCTTGGCCT	1084
Qy	1256	TTGGAGCCCTCACCATCACGCTGACCCCTGCAGGACACAGCCCTCGGTCCCTTACCTGA	1315
Db	1085	GCCTGGTGTACCGGTGGTGTCTCTATTCCAGAACAGGGTCCCGAGCTGTCTACCTCG	1144
Qy	1316	GTATCGTGGGCATTCTGGCCATCATCGCCTCTTTCTGAGTGGGCCAGGTGGCATCCCGT	1375
Db	1145	GCATCATCTGTCTTTGGCTTACATCGGGGACATTCATTTGGGCCAGTCTCTGTCCCT	1204
Qy	1376	TCATCTTCACTGGTGAATTTCCAGAAATCTCAGCGCGCGGTGCCTTTCATCATTTGCAG	1435
Db	1205	CGGTGGTCAGGACCGAGATCTTCTGCAGTCTCTCCCGGGGCGAGTTTCATGGTGGACG	1264
Qy	1436	GCACGTCACACTGGCTCTCCAACTTTGCTGTGGGCTCCTCTCCCATTCATTCCAGAAA	1495
Db	1265	GGGCAGTGCACCTGGGTCAACCAACTCATATAGGCTTCCCTGTCCCATCCATCCAGAGG	1324
Qy	1496	GTCTGGACACCTACTGTTTCTAGTCTTTGTGTACATTTGTATCACAGGTGCTATCTACC	1555
Db	1325	CCATCGGTGCCCTACAGTTTCATCTCTTTGCCGAACTGCTCTCTCACTGGGATTACA	1384
Qy	1556	TGTATTTTGTGCTGCTGAGACCAAAACAGAACCTATGCAGAAATCAGCCAGGCATTTT	1615
Db	1385	TCTACGTGTTATTCGGAGACCAAGGCGCAAAACATTTGTGGAGATAACCGCATTTTG	1444
Qy	1616	CCAAAGGAACAAAGCATACCCACAGAGAGAAATCGACTCAGCTGTCTACTGCTC	1675
Db	1445	CCAGAGAAACAGGGTGAAGCTTCCAGAGGAGAAAGAAACCATTTGATGCTGGGCTC	1504
Qy	1676	CTGCTTCTCTCTTCA	1693
Db	1505	CCACAGCCTCTCTGCCA	1522

Search completed: April 8, 2004, 14:27:32  
Job time : 9037 secs

## RESULT 15

CGGG 784



QY 864 CAACAGGCAAGAGCTGTGAAGCCCTCCAAAGCTTCTTGGTAAAGCAGCGTTTC 920  
 Db 541 CAACAGGCAAGAGCTGTGAAGCCCTCCAAAGCTTCTTGGTAAAGCAGCGTTTC 597

RESULT 13  
 AF490463 1296 bp mRNA linear ROD 28-MAY-2003  
 DEFINITION Mus musculus GLUT9a mRNA, complete cds.  
 ACCESSION AF490463  
 VERSION AF490463.1 GI:31088266  
 KEYWORDS  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 1296)  
 Carayannopoulos, M.O., Schlehn, A.L., Hoehn, A.L., Moley, J.F. and Moley, K.H.  
 Cloning and functional analysis of mouse GLUT9, a new embryonically expressed transporter  
 Unpublished  
 2 (bases 1 to 1296)  
 Carayannopoulos, M.O. and Moley, K.H.  
 Direct Submission  
 TITLE Submitted (06-MAR-2002) OB/Gyn, Washington University, 4911 Barnes Hospital Plaza, St. Louis, MO 63110, USA  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source 1..1296  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /chromosome="5"  
 /dev\_stage="embryonic day 7"  
 1..1296  
 /notes="member of the facilitative glucose transporter"  
 /codon\_start=1  
 /product="GLUT9a"  
 /protein\_id="AAP44162.1"  
 /db\_xref="GI:31088267"  
 /translation="WDSRELALASLMDTGGPGLSVGHQRTKWSFSLVVAALVG  
 AFSGFLYGNLSVNVNAPYKAFYNGTVRRHQPDIDPTLTLMSVTVSFAIYGG  
 LVGTLMKRMFLKRLSTLANNFPAISALLMACSLRAGTPEMLIVGVFMVDDG  
 IALSALPMYLNISPEKIRSLQGVTAIFIGVSGQLLGLPELLGRIMFVNSIFG  
 KAGIPDKPIYITLSTGGIETLAFISGLVIERLGRPLLGSGFLMALPFGTLTATL  
 TLQDAPWPYLSICILALIASPCSGGGIPFLITGEFTQSERPAFNIAGTVNWL  
 SNFAVLGIFPFFQKSDSYCFIVFATICAGATIFYFVLPEIKNRTHAEISQAFKRN  
 KACPPEVKADSAETBEKANSTPEPSSSTLDSYGONKIV"  
 627  
 /note="deletion of transmembrane regions 6 and 7"

misc\_feature  
 22.3%; Score 523.6; DB 10; Length 1296;  
 Best Local Similarity 65.8%; Pred. No. 9.6e-115; Indels 321; Gaps 1;  
 Matches 962; Conservative 0; Mismatches 179;

ORIGIN  
 211 AGGAGAAGAAAGACTGGTCCGCTCGCTCGCTGGGCTCCCTCGGGGGCGCTTCGGC 270  
 Db 82 AGGAGACAAAGAAATGGTCTCTCGCTCGCTGGGAGCTTCGTTGGAGCGCTTCGG 141  
 QY 271 TCCTCTCTCTCTAGCTTACAACTGTGGTGGTGAATGCCCAACCCGCTACATCAAG 330  
 Db 142 TCCTCTCTCTCTAGCTTACAACTGTGGTGGTGAATGCCCAACCCGCTATATCAAG 201  
 QY 331 GCCTTTTCAATGAGTCAATGGGAAGAGGATGAGCTCCCAATAGACCAGACACTCTG 390  
 Db 202 GCCTTTTCAATGAGTCAATGGGAAGAGGATGAGCTCCCAATAGACCAGACACTCTG 261  
 QY 391 ACTCTCTCTCTGTGTGACTGTGTCCATATTCGCAATCGGTGACTTGTGGGAGCGTTA 450  
 Db 262 ACTCTCTCTGTGTGACTGTGTCCATATTCGCAATCGGTGACTTGTGGGAGCGCTG 321

QY 451 ATTGTGAAGATGATTGGAAAGTTCTTGGGAGGAAGCACAACCTTTGCTGGCCTATATGGG 510  
 Db 322 ATGTGTGAAGATGATTGGAAAGTTCTTGGGAGGAAGTCCACATTTGCTGGTCAACAATGGA 381  
 QY 511 TTTCAATTTCTGTGCAATTTGCTGATGGCTGTCTCGTCCAGGACGAGAGCTTTGAAATG 570  
 Db 382 TTTGCCATTTCCCGCAGCGTTGCTGATGGCATGTCTACTCCGGGACGAACTTTTGAGATG 441  
 QY 571 CTCATTGTGGAGCGCTTCATCATGGGATAGATAGGAGGCTCGCCCTCAGTGTCTCCGCC 630  
 Db 442 CTCATTGTGGAGCGCTTCATCATGGGATAGATAGGAGGCTTCATCTCAGCGCGCTCCGCC 501  
 QY 631 ATGTACCTCAGTGAGATCTCACCAAGGAGATCCGTGGCTCTCTCTGGGGAGGAGTCACTGCC 690  
 Db 502 ATGTATCTACGAGATCTCACCAAGGAGATCCGGGGCTCTCTCTGGGGAGGAGTCACTGCC 561  
 QY 691 ATCTTTATCTCATTTGGCTGTCTCAGTGGGAGCTTCTGGGCTGCTGCCAGAGCTGTGGGA 750  
 Db 562 ATCTTCATCTCATTTGGCTGTCTCAGTGGGAGCTTCTGGGCTGCTGCCAGAGCTGTGGGA 621  
 QY 751 AAGGAGAGTACCTGGCCATACCTGTTTGGAGTATTGTGGTCCCTGCCGTGTGCCAGCTG 810  
 Db 622 AGG----- 624  
 QY 811 CTGAGCGCTTCCCTTTCTCCGGGACAGCCACGCTACCTGCTCTTGGAGAGCACAACAGAG 870  
 Db 625 ----- 624  
 QY 871 GCAAGAGCTGTGAAGAGCTTCCAAACGTTCTTGGGTAAAGCAGACGTTTCCCAAGAGGTA 930  
 Db 625 ----- 624  
 QY 931 GAGGAGGTCCTGGCTGAGAGCCACGTCAGAGGAGCATCCGCTGCTGCTCGTGTGGAG 990  
 Db 625 ----- 624  
 QY 991 CTGCTGAGAGCTCCCTAGCTCCGCTGGCAGGTGGTTCACCGTGAATGTTCACCATGGCGCTGC 1050  
 Db 625 ----- 624  
 QY 1051 TACAGAGCTCTGGCGCTCAATGCAATTTGGTTCTATACCAACAGAGCATTTTGGAAAGCT 1110  
 Db 625 ----- ATCTGGTTCTACCAACAGAGCATTTTGGAAAGCT 660  
 QY 1111 GGGATCCCTCGGCAAGATCCCATAGTCACTTCACTAGTACAGGGGGCATTCGAGACTTGG 1170  
 Db 661 GGGATCCCTCAGGACAAGATCCCATATCATCCCTGAGCAGCGGAGGAATGAAACACTG 720  
 QY 1171 GCTGCGGCTTCTCTGTTTGGTCAATGAGCAGCTGGGAGCGGAGACCCCTCTCTCATTTGT 1230  
 Db 721 GCTGCTATCTTCTCTGGCTTGGTCAATGAGCGCTTAGGAAGGAGAGCTCTCTCTCATTTGT 780  
 QY 1231 GGGCTTGGGCTCATGSGGCTCTTCTTTGGGAGCCCTCACCATCACGCTGACCCCTGACAGGAC 1290  
 Db 781 GGGCTTGGGCTCATGSGGCTCTTCTTTGGGAGCCCTCACCGGACACTGACCCCTGACAGGAC 840  
 QY 1291 CACGCCCCCTGGGTCCCTTACCTGAGTATCGTGGGATTTCTGGCCATCATGCGCTCTTTC 1350  
 Db 841 CAAAGTCCCTGGGTCCCTTACCTCAGCATTTGTGTCATCTTGGCCATCATTTGCTGCTTC 900  
 QY 1351 TGGAGTGGGCGAGGTGGCATCCGTTCACTTCACTGGTGGTGTCTTCCAGGAACTCTAG 1410  
 Db 901 TGCAGCGGTCCAGGTGGGATCCCATTCATCTCAGCGGAGAGTTCCTTCAACAGTCAAG 960  
 QY 1411 CGGCGGCTGCTCTCATTCATTCATTGAGGACCGCTCAACTGGCTCTTCCAACTTTTGTGTGG 1470  
 Db 961 CGACCGGAGCGCTTCATGATCGGAGGACAGTCAACTGGCTCTCAAACTTCGCGGTAGGG 1020  
 QY 1471 CTCCTCTTCCCATTCATTCAGAAAAGTCTGGACCACTACTGTCTTCTAGTCTTGTGTACA 1530  
 Db 1021 CTCCTTTTCCCTTTCATCCAGAAAAGCTGGAGTCTCTACTGTCTTCTCTCTTCTTCTGCGACA 1080  
 QY 1531 ATTGTATACAGGTGCTATCTACCTGTATTTTGTGCTGCTGAGACCAAAACAGAAC 1590







Db 111121 GTCAGGAGATTGAGACCACTCTGGCTAACATGGTGAAACTCCCTCTCTACTTAAATAACA 111062  
 QY 2150 AAAATTAGCTGAGCATGGTGGCGGCGCTGTAGTCCAGCTACTTGGGAGGCTGAGGC 2209  
 Db 111061 AAAATTAGCTGAGCATGGTGGCGGCGCTGTAGTCCAGCTACTTGGGAGGCTGAGGC 111002

QY 2210 AGGAGATGGCGTGAACCCAGGAGCGGAGTTGCGAGTGACCGAGATCACACACACCA 2269  
 Db 111001 AGGAGATGGCGTGAACCCAGGAGCGGAGTTGCGAGTGACCGAGATCACACACACCA 110942

QY 2270 CTCAGCTGGTGTACAGACGACGACTCCGTCTCAAAAAAATAAAAAA 2326  
 Db 110941 CTCAGCTGGTGTACAGACGACGACTCCGTCTCAAAAAAATAAAAAA 110885

RESULT 10  
 AC098976/c 165221 bp DNA linear PRI 01-MAR-2002  
 LOCATION Homo sapiens BAC clone RP11-751L19 from 4, complete sequence.  
 DEFINITION AC098976 AC068403  
 ACCESSION AC098976.2 GI:18042509  
 VERSION HTG.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Sulston, J.B. and Waterston, R.  
 1 (bases 1 to 165221)  
 Toward a complete human genome sequence  
 Genome Res. 8 (11), 1097-1108 (1998)  
 99063792  
 9847074

REFERENCE 2 (bases 1 to 165221)  
 AUTHORS Levy, A. and Cotton, M.  
 TITLE The sequence of Homo sapiens BAC clone RP11-751L19  
 JOURNAL Unpublished (2001)  
 MEDLINE 3 (bases 1 to 165221)  
 REFERENCE 3 (bases 1 to 165221)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-NOV-2001) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 4 (bases 1 to 165221)  
 REFERENCE 4 (bases 1 to 165221)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-JAN-2002) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 5 (bases 1 to 165221)  
 REFERENCE 5 (bases 1 to 165221)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-MAR-2002) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On Jan 3, 2002 this sequence version replaced gi:16756380.  
 ----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc>  
 Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
 ----- Summary Statistics  
 Center project name: H\_NH0751L19  
 Drafting Center: WIBR  
 -----

NOTICE: This sequence may not represent the entire insert of this  
 clone. It may be shorter because we only sequence overlapping  
 clone sections once, or longer because we provide a small overlap  
 between neighboring data submissions.  
 This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate  
 chemistry, or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence  
 from more than one subclone; and the assembly was confirmed by  
 restriction digest.  
 MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. John D.  
 McPherson, Department of Genetics, Washington University, St. Louis  
 MO. For additional information about the map position of this  
 sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:  
 The RPCI-11 human BAC library was made from the blood of one male  
 donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frenken, E.,  
 Tatenio, M., Catanese, J., and de Jong, P.J. (1998) An improved  
 approach for construction of bacterial artificial chromosome  
 libraries. Genomics 51:1-8. The clone may be obtained either from  
 Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong  
 and coworkers at <http://www.chori.org>  
 VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:  
 Actual start of this clone is at base position 1 of RP11-751L19;  
 actual end is at base position 165221 of RP11-751L19.

Data from AC067775 was used to finish this clone, AC098976.  
 The sequence of AC068403 has been incorporated into AC098976.

FEATURES	Source	Location/Qualifiers
	1..165221	/organism="Homo sapiens"
		/mol_type="Genomic DNA"
		/db_xref="taxon:9606"
		/chromosome="4"
		/map="4"
		/clone="RP11-751L19"
		/clone_lib="RPCI-11"
misc_feature	1..154	/note="match to EST B5062468 (NID:38407118)"
misc_feature	161..427	/note="similar to Homo sapiens EST BF878566 (NID:912268696)"
repeat_region	173..290	/rpt_family="L2"
repeat_region	915..1201	/rpt_family="Alu"
repeat_region	1203..1425	/rpt_family="MER2_type"
repeat_region	1638..1952	/rpt_family="Alu"
repeat_region	1823..1954	/rpt_family="(A)n"
repeat_region	3476..3772	/rpt_family="Alu"
repeat_region	3773..3909	/rpt_family="ERV1"
misc_feature	3788..3793	/note="similar to Homo sapiens EST BF171573 (NID:913437787)"
repeat_region	3910..4193	/rpt_family="MaLR"
repeat_region	4194..4507	/rpt_family="Alu"
misc_feature	4194..4208	/note="similar to Homo sapiens EST AW451324 (NID:96992100)"
misc_feature	4194..4206	/note="similar to Homo sapiens EST AI371376 (NID:94150129)
misc_feature	4194..4206	/note="similar to Homo sapiens EST AI740932 (NID:95109220)
repeat_region	4508..4704	/rpt_family="MaLR"





VERSION AXI79745.1 GI:15132109  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1  
 AUTHORS Baughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R.,  
 Lai, P., Hillman, J.L., Azimzai, Y., Yue, H., Nguyen, D.B., Yao, M.G.,  
 Gandhi, A.R., Tang, Y.T. and Khan, F.A.  
 TITLE Transporters and ion channels  
 JOURNAL Patent: WO 0146258-A 33 28-JUN-2001;  
 FEATURES Incyte Genomics, Inc. (US)  
 source Location/Qualifiers  
 1. 2216  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 6618083CB1"  
 ORIGIN  
 Query Match 48.7%; Score 1141; DB 6; Length 2216;  
 Best Local Similarity 99.1%; Pred. No. 5.4e-263;  
 Matches 1147; Conservative 0; Mismatches 10; Indels 0; Gaps 0;  
 217 AGAAGGACTGGTCTCTCGCTCCCTCGTGGCCCTCGCTCGCGGGGCGCTTCGGTCCCTCC 276  
 157 AAATTGGACTGGTCTCTCTCGCTCCCTCGTGGCCCTCGCTCGCGGGGCGCTTCGGTCCCTCC 216  
 277 TTCCTCTACGGCTACAACTGTTCGGTGGTGAATGCCCCCAACCCCGGTACATCAAGGCCCTTT 336  
 217 TTCCTCTACGGCTACAACTGTTCGGTGGTGAATGCCCCCAACCCCGGTACATCAAGGCCCTTT 276  
 337 TACAATGAGTCATGGGAAAGAGGCATGGAGCGTCCCAATAGACCCAGACACTCTGACTCTG 396  
 277 TACAATGAGTCATGGGAAAGAGGCATGGAGCGTCCCAATAGACCCAGACACTCTGACTCTG 336  
 397 CTCTGGTCTGTGACTGTCCATATTTCGCCATCGGTGGACTTGTGGGACGTTAATTTGTG 456  
 337 CTCTGGTCTGTGACTGTTCATTTTCGCCATCGGTGGACTTGTGGGACGTTAATTTGTG 396  
 457 AAGATGATGGAAAGTTCTTGGGAGGAAGCACTTTGCTGCCAATAATAGGTTTGA 516  
 397 AAGATGATGGAAAGTTCTTGGGAGGAAGCACTTTGCTGCCAATAATAGGTTTGA 456  
 517 ATTTCTGCTGATTCGATGGCCCTGCTGCTCCAGGACGAGACCTTTGAAATGCTCAT 576  
 457 ATTTCTGCTGATTCGATGGCCCTGCTGCTCCAGGACGAGACCTTTGAAATGCTCAT 516  
 577 GTGGGACGCTTCATATGGGCATAGATGAGGCGTCGCCCTCAGTGTGCTCCCATGAC 636  
 517 GTGGGACGCTTCATATGGGCATAGATGAGGCGTCGCCCTCAGTGTGCTCCCATGAC 576  
 637 CTCAGTGAGATCTCACCCAAAGAGATCCGTTGCTCTCTGGGGCAGGTGACTGCCATCTTT 696  
 577 CTCAGTGAGATCTCACCCAAAGAGATCCGTTGCTCTCTGGGGCAGGTGACTGCCATCTTT 636  
 697 ATCTGCAATGGCGTGTTCATCTGGGACAGCTTTCGGGCGCTGCCGAGAGCTGTGGGAAAGAG 756  
 637 ATCTGCAATGGCGTGTTCATCTGGGACAGCTTTCGGGCGCTGCCGAGAGCTGTGGGAAAGAG 696  
 757 AGTACCTGGCCATACCTGTTTGGAGTGATTTGGTCCCTGCGCTGCTCCACTGCTGAGC 816  
 697 AGTACCTGGCCATACCTGTTTGGAGTGATTTGGTCCCTGCGCTGCTCCACTGCTGAGC 756  
 817 CTTCCCTTTCTCCCGACACCCACGCTACCTGCTCTTGGGAAAGCAACAAGGACAGA 876  
 757 CTTCCCTTTCTCCCGACACCCACGCTACCTGCTCTTGGGAAAGCAACAAGGACAGA 816  
 877 GCTGTGAAGCCTTCCAAACGTTCTTTGGGTAAAGCAGACGTTTCCCAAGAGGTAGAGAG 936  
 817 GCTGTGAAGCCTTCCAAACGTTCTTTGGGTAAAGCAGACGTTTCCCAAGAGGTAGAGAG 876

Db 1657 TCAGCTGTCACTGATGTT 1674

RESULT 6  
LOCUS AR339083 2842 bp DNA linear PAT 17-AUG-2003  
DEFINITION Sequence 574 from patent US 6569662.  
ACCESSION AR339083  
VERSION AR339083.1 GI:33725940  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 2842)  
AUTHORS Tang, Y.T., Zhou, P. and Drmanac, R.T.  
TITLE Nucleic acids and polypeptides  
JOURNAL Patent: US 6569662 A 574 27-MAY-2003;  
FEATURES Location/Qualifiers  
source 1..2842  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN

Query Match 56.2%; Score 1317.8; DB 6; Length 2842;  
Best Local Similarity 89.9%; Pred. No. 1.9e-305;  
Matches 1504; Conservative 0; Mismatches 7; Indels 162; Gaps 2;

QY 2 CGACCCACCGCTCCGCGCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGT 61  
Db 40 CGACCCACCGGTGC-GCCCTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGT 98  
QY 62 CACTGAGACCCATGGCAAGAAACAAATAGGAATTCGAAGAACTGGGCTAGTTCCTCC 121  
Db 99 CACTGAGACCCATGGCAAGAAACAAATAGGAATTCGAAGAACTGGGCTAGTTCCTCC 158  
QY 122 TCACAGATGACACACGCCACGCCGCTCCAGGCGCAGAGGAGGCACTGTGAGAGTG 181  
Db 159 TCACAGATGACACACGCCACGCCGCTCCAGGCGCAGAGGAGGCACTGTGAGAGTG 218  
QY 182 ACCACCTGAGGAGTGGGGTCCAGGTGGAAGAGAGAGAGAGAGAGTGTCTGCTGCC 241  
Db 219 ACCACCTGAGGAGTGGGGTCCAGGTGGAAGAGAGAGAGAGAGTGTCTGCTGCC 278  
QY 242 TCGTGGCTCCCTCCGCGGGCGCTTCGGCTCCCTCTCTCTACCGCTACAACTGTCGG 301  
Db 279 TCGTGGCTCCCTCCGCGGGCGCTTCGGCTCCCTCTCTCTACCGCTACAACTGTCGG 338  
QY 302 TGTGTAATGCCCGCCACCCCGTACATCAAGGCCCTTTTACAATGAGTCAATGGAAAGAGC 361  
Db 339 TGTGTAATGCCCGCCACCCCGCTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCC 357  
QY 362 ATGACGCTCCATAGACCCAGACACTCTGACTCTGCTGCTGCTGCTGCTGCTGCTGCT 421  
Db 358 ----- 357  
QY 422 TCGCCATCGGTGAGCTTTGGGAGCGTTAATTTGAAGATGATTGGAAGGTTCTTGGGA 481  
Db 358 ----- 357  
QY 482 GGAAGACACTTTGCTGGGCAATATAGGTTTGGCAATTTCTGCTGCAATGCTGATGGCCT 541  
Db 358 GGAAGACACTTTGCTGGGCAATATAGGTTTGGCAATTTCTGCTGCAATGCTGATGGCCT 417  
QY 542 GCTCGCTCCAGGCGAGGAGCTTTGAAATGCTCATTTGCGAGCGCTTCATCATGCGGCATAG 601  
Db 418 GCTCGCTCCAGGCGAGGAGCTTTGAAATGCTCATCGTGGAGCGCTTCATCATGCGGCATAG 477  
QY 602 ATGAGGCGCTCGCCCTCAGTGTGCTCCCATGCTACCTCAGTCAGATCTCAACCCAGAGGA 661  
Db 478 ATGAGGCGCTCGCCCTCAGTGTGCTCCCATGCTACCTTAGTGAGATCTCAACCCAGAGGA 537  
QY 662 TCGTGGCTCTCTGGGCGAGGTGACTGGCACTCTTTATCTGCAATTTGAGCTGTCTGATGGC 721

Db 538 TCCGTGGCTCTCTGGGCGAGGTGACTGCCATCTTTATCTGCAATGGCGTGTTCACCTGGGC 597  
QY 722 AGCTTCTGGGCTGCCCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAG 781  
Db 598 AGCTTCTGGGCTGCCCCGAGCTGCTGGGAAAGAGAGTACCTGGCCATACCTGTTTGGAG 657  
QY 782 TGATTTGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 841  
Db 658 TGATTTGCTGCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717  
QY 842 GCTACCTGCTGCTGGAGAGACAAACGAGGCGAGAGCTGTGAAAGCTTCCAAAGCTTCT 901  
Db 718 GCTACCTGCTGCTGGAGAGACAAACGAGGCGAGAGCTGTGAAAGCTTCCAAAGCTTCT 777  
QY 902 TGGGTAAGCAGACAGTTTCCCAAGAGGTAGAGGAGTCTCGCTGAGAGCAGCAGTGCAGA 961  
Db 778 TGGGTAAGCAGACAGTTTCCCAAGAGGTAGAGAGTCTCGCTGAGAGCAGCAGTGCAGA 837  
QY 962 GGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1021  
Db 838 GGAGCATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897  
QY 1022 TGGTCAACCTGATGTACCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1081  
Db 898 TGGTCAACCTGATGTACCAATGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957  
QY 1082 TCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGCGCAAGATCCCATACGTC 1141  
Db 958 TCTATACCAACAGCATCTTTGGAAAGCTGGGATCCCTCCGCGCAAGATCCCATACGTC 1017  
QY 1142 CTTGAGTACAGGGGGCATCGAGACTTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1201  
Db 1018 CTTGAGTACAGGGGGCATCGAGACTTTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1077  
QY 1202 ACTGAG 1261  
Db 1078 ACTGAG 1137  
QY 1262 CCTCACCACATCAGCTGACCTGAGAGACACGCGCCCTGAGGTCCTTACCTACCTGAGTATCG 1321  
Db 1138 CCTCACCACATCAGCTGACCTGAGAGACACGCGCCCTGAGGTCCTTACCTGAGTATCG 1197  
QY 1322 TGGGCAATCTGCGCATCATCGCTCTTTCTGAGTGGGCGCAGGTGGCATCCCGTTCATCT 1381  
Db 1198 TGGGCAATCTGCGCATCATCGCTCTTTCTGAGTGGGCGCAGGTGGCATCCCGTTCATCT 1257  
QY 1382 TGACTGGTGGTCTTCCAGCAATCTCAGCGCGCGCTGCTTCAATTCATTCAGAGGACCG 1441  
Db 1258 TGACTGGTGGTCTTCCAGCAATCTCAGCGCGCGCTGCTTCAATTCATTCAGAGGACCG 1317  
QY 1442 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCTCTTCCCAATTCATTCAGAGGAGTCTGG 1501  
Db 1318 TCAACTGGCTCTCCAACTTTGCTGTTGGGCTCTCTTCCCAATTCATTCAGAGGAGTCTGG 1377  
QY 1502 ACACCTACTGTTTCTAGTCTTTGCTACAAATTTGATACAGGTGCTATCTACCTGATTT 1561  
Db 1378 ACACCTACTGTTTCTAGTCTTTGCTACAAATTTGATACAGGTGCTATCTACCTGATTT 1437  
QY 1562 TTGCTGCTGCTGAGACCAAAACAGAACTATGAGAAATCAGGCGAGCATTTTCCAAA 1621  
Db 1438 TTGCTGCTGCTGAGACCAAAACAGAACTATGAGAAATCAGGCGAGCATTTTCCAAA 1497  
QY 1622 GGAACAAAGCATACCCACAGAGAGAAATCGACTCAGCTGTCACTGATGCT 1674  
Db 1498 GGAACAAAGCATACCCACAGAGAGAAATCGACTCAGCTGTCACTGATGCT 1550

RESULT 7  
AX179745  
LOCUS AX179745  
DEFINITION Sequence 33 from Patent WO0146258.  
ACCESSION AX179745

AX179745 2216 bp DNA linear PAT 06-AUG-2001



Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Larić, P., Legaspi, R.,  
Maduro, Q. L., Mastello, C., Maskeri, B., Mastrian, S. D., McCloskey, J. C.,  
McDowell, J., Pearson, R., Stantripop, S., Thomas, P. J., Touchman, J. W.,  
Tsurgeon, C., Vogt, J. L., Walker, M. A., Wetherby, K. D., Wiggins, L.,  
Young, A., Zhang, L.-H. and Green, E. D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
 Series: IRAL Plate: 15 Row: b Column: 17.

FEATURES	SOURCE
----------	--------

```

1. 1885
/organism="Homo sapiens"
/mol_type="RNA"
/db_xref="taxon:9606"
/clone="MGC:11313 IMAGE:3949549"
/tissue_type="Ovary, adenocarcinoma"
/clone_lib="NIH_MGC_9"
/lab_host="DR10B-R"
/notes="Vector: DOT37"

```

gene

```

/ gene="SUCC2A3"
/ note="synonyms: GLUTX, Glut9"
/ db_xref="LocusID:56606"
/ db_xref="MIM:606142"
160: .1695

```

CDs

```

/codon_start=1
/product="Slc2A9 protein"
/protein_id="AAH18897.1"
/db_xref="GI:17511906"
/db_xref="Locustid:56606"
/translation="MKLSKKRQDEEDSSAKKLDNWSLLVASIAGAFSGSLFYIG
NLSVNAPYPIKATYNSWERRHGRIPDPTLTLLNWSVTSFATGGVGLTIVLQMI
KGVLRGRKTHLLNANGFAISAAIIMACSLQAGAFEMLIAGVFIMGIDGVALSVPL
SLSPKEIKRISLQGVNTAFICIGVFTQGLLGLPGLKGSWTFPFGVIVPAAVQGL
SIPPLDPSRYLLLEKHNARAKAFQFLCKADGSEVLEAESRORSIRLRSLV
ELLARAPYRWQVTVVIMNACVQLGNAIYNTNSIFGKAGIPLAKIIPVTLTSGI
ELTAUAFSGSLVIEHLGRPLLIIGLFGELMGLFEGTLTITLTDHPHWPVYLSIVGILA
ITASFCSGGITGPIFLUTEVFPQOSQRPAAFIAGTVMNLSENVAFLGLFPFIQSLDTY
CELVAFATICTAIYLYFVLPEKNTNRYAESQAFSKRKAIPPEEKISDAVTDGKIN
GRF"
241. 1614
/notes="sugar tr; region: (and other) transporter"
/db_xref="CDD:fam0083"

```

## ORIGIN

Query Match	61.7%	Score 1415.2;	DB 9;	Length 1885;
Best Local Similarity	99.5%;	Pred. No. 0;		
Matches 1450: Conservative	0;	Mismatches 8;	Indels 0;	Gaps 0;

217	AGAAAGGACTGGTCTGCTCGCTCTCGTGGGCTCCCTCGGGGGCGCTTCGGGTCTCC	276
217	AAATTGGACTGGTCTGCTCGCTCTCGTGGGCTCCCTCGGGGGCGCTTCGGGTCTCC	276
277	TTCCCTCTACGGGTCAAACTGTGCGTGGTGAATGCCCCACCCCGTACATCAAGGCCCTTT	336
277	TTCCCTCTACGGGTCAAACTGTGCGTGGTGAATGCCCCACCCCGTACATCAAGGCCCTTT	336
337	TACAAATGAGTCATGGGAAGAGGCATGAGACCTGCAATAGACCCAGACACTCTGACTCTG	396
337	TACAAATGAGTCATGGGAAGAGGCATGAGACCTGCAATAGACCCAGACACTCTGACTCTG	396
397	CTCTGGCTGTGACTGTGTCATATTTCGGCATTCGGTGGACTTGTGGGACGCTTAATCTG	456
397	CTCTGGCTGTGACTGTGTCATATTTCGGCATTCGGTGGACTTGTGGGACATTAATCTG	456
457	AAGATGATTTGGAAAGGTTCTTGGAGGAAGCACTTTGTCGGCCAAATAATGGGTTGCA	516
457	AAGATGATTTGGAAAGGTTCTTGGAGGAAGCACTTTGTCGGCCAAATAATGGGTTGCA	516
517	ATTTCTGCTGCAATGCTGATGGCGCTGCTCGCTCCAGGCAAGAGCCCTTTGAAATGCTCAT	576
517	ATTTCTGCTGCAATGCTGATGGCGCTGCTCGCTCCAGGCAAGAGCCCTTTGAAATGCTCAT	576





1441	Db	GTCAACTGGCTCTCCAACTTTGCTGTGGGCTCCTCTCCCAATCAATTCAGAAAAGTCTG	1500
1501	Qy	GACCACTACTGTTTCCCTAGTCTTTTGTCACAAATTTGTATCAGAGTGCTATCTACCTGTAT	1560
1501	Db		
1501	Db	GACCACTACTGTTTCCCTAGTCTTTTGTCACAAATTTGTATCAGAGTGCTATCTACCTGTAT	1560
1561	Qy	TTTGTGCTGCTGGAGCCAAAACAGAACTTATGCAGAAATCAGCCAGGCATTTTCCAAA	1620
1561	Db		
1561	Db	TTTGTGCTGCTGGAGCCAAAACAGAACTTATGCAGAAATCAGCCAGGCATTTTCCAAA	1620
1621	Qy	AGGAACAAAGCATACCCACCAGAGAAGAAAAATCGAATCAGCTGTCACTGATGCTCCTGCT	1680
1621	Db		
1621	Db	AGGAACAAAGCATACCCACCAGAGAAGAAAAATCGAATCAGCTGTCACTGATGCTCCTGCT	1680
1681	Qy	TTCTTCTCCTTTCACTACTCCGAATACAGCCTGGATTCAAGCTGGCGGCAACACACCGGC	1740
1681	Db		
1681	Db	TTCTTCTCCTTTCACTACTCCGAATACAGCCTGGATTCAAGCTGGCGGCAACACACCGGC	1740
1741	Qy	ACCAAAAAGAAACACCCATTGTAAACGGTCACTGGCTATTTCCTCAACCTGGGAATGACCT	1800
1741	Db		
1741	Db	ACCAAAAAGAAACACCCATTGTAAACGGTCACTGGCTATTTCCTCAACCTGGGAATGACCT	1800
1801	Qy	TCCCTATPTCTTCTCTCGGAAACACCAAGTCATGATGCAGACAGAGCTTGGATTT	1860
1801	Db		
1801	Db	TCCCTATPTCTTCTCTCGGAAACACCAAGTCATGATGCAGACAGAGCTTGGATTT	1860
1861	Qy	TGGAGACATGGTTTGAATTTCCAGTCATTCACTCTTTTATTTCAGCAATATTTAAACAAGT	1920
1861	Db		
1861	Db	TGGAGACATGGTTTGAATTTCCAGTCATTCACTCTTTTATTTCAGCAATATTTAAACAAGT	1920
1921	Qy	ACTGACATGTCCCATATGTGTGTTTACCACACTGGTTATACAATGGGAGGAGAGAGAG	1980
1921	Db		
1921	Db	ACTGACATGTCCCATATGTGTGTTTACCACACTGGTTATACAATGGGAGGAGAGAGAG	1980
1981	Qy	AGAGAGAGAGAGAGAGATGCTATTCTTAAAGCTTTGAAGTCTAGGCTGTGCACGGTGCTC	2040
1981	Db		
1981	Db	AGAGAGAGAGAGAGAGATGCTATTCTTAAAGCTTTGAAGTCTAGGCTGTGCACGGTGCTC	2040
2041	Qy	ACGGCTCTAATCCACGACCTTTTGGGAGGCCGAGGTGGGTGGATCGTGAGCTCAGGAGATT	2100
2041	Db		
2041	Db	ACGGCTCTAATCCACGACCTTTTGGGAGGCCGAGGTGGGTGGATCGTGAGCTCAGGAGATT	2100
2101	Qy	GAGACCATCTCTGGCTAACATGTTGMAACTCCCTCTCTACTAAAAATACAAAAATTTAGCT	2160
2101	Db		
2101	Db	GAGACCATCTCTGGCTAACATGTTGMAACTCCCTCTCTACTAAAAATACAAAAATTTAGCT	2160
2161	Qy	GAGCATGCTGGCGGGCGCTGTAGTCCACAGCTACTTTGGGAGGCTGAGCGAGGAGATGCG	2220
2161	Db		
2161	Db	GAGCATGCTGGCGGGCGCTGTAGTCCACAGCTACTTTGGGAGGCTGAGCGAGGAGATGCG	2220
2221	Qy	GTGAACCCAGAGAGCGGAGGATTGCGATGAGCGGAGATCACACACACACTCCAGCGCTGG	2280
2221	Db		
2221	Db	GTGAACCCAGAGAGCGGAGGATTGCGATGAGCGGAGATCACACACACACTCCAGCGCTGG	2280
2281	Qy	GTGACAGAGCGACACTCCGTCCTCAAAAAAATAAAAAAATAAAAAAAGGCGCGC	2340
2281	Db		
2281	Db	GTGACAGAGCGACACTCCGTCCTCAAAAAAATAAAAAAATAAAAAAAGGCGCGC	2340
2341	Qy	CGC 2343	
2341	Db		
2341	Db	CGC 2343	

[illegible]

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

1

REFERENCE  
AUTHORS  
Bruns, C.M., Sellinger, J.J., Tribouley, C.M., Azinai, Y., Ding, L.,  
Hafalia, A., Khan, P.A., Yao, M.G., Yue, H., Au-Young, J., Baughn, M.R.,  
Borowsky, M.L., Burford, N., Elliott, V.S., Gandhi, A.R., Green, B.D.,  
Kearney, J., Lai, P., Lee, E.A., Lu, D.A.M., Lu, Y., Tang, Y.T., Yang, J.,  
Griffin, J.A., Hillman, J.L., Nguyen, D.B., Policky, J.L., Ramkumar, J.,  
Raumann, B.E., Sanjanwalla, M.S., Thornton, M., Wallia, N.K.  
TITLE  
JOURNAL  
Patent: WO 0192304-A 52 06-DEC-2001;  
Incyte Genomics, Inc. (US)  
FEATURES  
source  
Location/Qualifiers  
1. .2032  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="Incyte ID No: 7476643CB1"

ORIGIN

Query Match 70.5%; Score 1651; DB 6; Length 2032;  
Best Local Similarity 99.7%; Pred. No. 0;  
Matches 1654; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 17 GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG 76  
Db 1 GCCTTGGCAGAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGGTCACTGAGACCCATGG 50

QY 77 CAAGGAAACAAATAGGAATCCAAAGAACTGGGCTAGTTCCTCTCA CAGATGACCA 136  
Db 61 CAAGGAAACAAATAGGAATCCAAAGAACTGGGCTAGTTCCTCTCA CAGATGACCA 120

QY 137 GCCACCGGGCTCCAGGCCAGGGAGGCACCTGCTGGAGTGTGACCACTGAGAGTG 196  
Db 121 GCCACCGGGCTCCAGGCCAGGGAGGCACCTGCTGGAGTGTGACCACTGAGAGTG 180

QY 197 GGGTGCAGGTGGAGGAGAAAGAGACTGGTCTGCTCGCTCTGCTGGGCTCCCTCG 256  
Db 181 GGGTGCAGGTGGAGGAGAAAGAGACTGGTCTGCTCGCTCTGCTGGGCTCCCTCG 240

QY 257 CGGGGCGCTTGGGCTCTCTCTCTAAGGCTACAACTGCTGGTGGTGAATGCCCA 316  
Db 241 CGGGGCGCTTGGGCTCTCTCTCTCTAAGGCTACAACTGCTGGTGGTGAATGCCCA 300

QY 317 CCCGTGACATCAAGGCCCTTTTACAATGAGTCATGGGAAAGGACATGACGCTCAAATAG 376  
Db 301 CCCGTGACATCAAGGCCCTTTTACAATGAGTCATGGGAAAGGACATGACGCTCAAATAG 360

QY 377 ACCGAGACACTGACACTGCTCTGGTCTGTGACTGTGTCCATATTGCGCAATCGGTGAC 436  
Db 361 ACCGAGACACTGACACTGCTCTGGTCTGTGACTGTGTCCATATTGCGCAATCGGTGAC 420

QY 437 TTGTGGGACGTTAATTCGTAAGATGATGGAAAGGTTCTTGGGAGGAAGACACTTGC 496  
Db 421 TTGTGGGACGTTAATTCGTAAGATGATGGAAAGGTTCTTGGGAGGAAGACACTTGC 480

QY 497 TGCCCAATTAATGGGTTTGCATTTCTGCTGCATTTGCTGATGGCTGCTGCTGCCAGGAG 556  
Db 481 TGCCCAATTAATGGGTTTGCATTTCTGCTGCATTTGCTGATGGCTGCTGCTGCCAGGAG 540

QY 557 GAGCCTTTGAATGCTCAATGAGGAGCTTCATCATGGGCAATGATGAGGCGTGCACC 616  
Db 541 GAGCCTTTGAATGCTCAATGAGGAGCTTCATCATGGGCAATGATGAGGCGTGCACC 600

QY 617 TCAGTGTGCTCCCAATGACCTCAGTGAGATCTCACCCAGGAGATCCGTGGCTCTCTGG 676  
Db 601 TCAGTGTGCTCCCAATGACCTCAGTGAGATCTCACCCAGGAGATCCGTGGCTCTCTGG 660

QY 677 GGCAGGTGACGCCATCTTTATCTGCAATGGCGTGTCTCACTGGGAGGCTTTGGGCGCTGC 736  
Db 661 GGCAGGTGACGCCATCTTTATCTGCAATGGCGTGTCTCACTGGGAGGCTTTGGGCGCTGC 720

QY 737 CCGAGCTGCTGGAAAGGAGAGTAGTACCTGGGCCATACCTGTTGGAGTGAATTTGGTGCCTG 796

Db		2041	AGCCCTGTAATCCCGACACTTTTGGGAGGCCGAGGTGGGTGGATCGTGAGSTCAGAGATT	2100
QY		2101	GAGCACATTCCTGGCTAAACATGGTGAACATCCCCTCTCTACTAAAAATACAAAATAATTAGCT	2160
Db		2101	GAGCACATTCCTGGCTAACATGGTGAACATCCCCTCTCTACTAAAAATACAAAATAATTAGCT	2160
QY		2161	GAGCATGGTGGCGGGCCCTGTATCCAGACTACTTTGGAGGCTTGAAGCAGAGAAATGGC	2220
Db		2161	GAGCATGGTGGCGGGCCCTGTATCCAGACTACTTTGGAGGCTTGAAGCAGAGAAATGGC	2220
QY		2221	GTGAACCCAGGAGCGGAGCTTGCAGTGAGCCGAGATCACACCACCACTCCAGAGCTGG	2280
Db		2221	GTGAACCCAGGAGCGGAGCTTGCAGTGAGCCGAGATCACACCACCACTCCAGAGCTGG	2280
QY		2281	GTGCAGAGCCAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGC	2340
Db		2281	GTGCAGAGCCAGACTCCGTCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAGGCGCGC	2340
QY		2341	CGC 2343	
Db		2341	CGC 2343	
RESULT 2				
LOCUS		AR184379		
DEFINITION		Sequence 1 from patent US 6346374.	DNA	linear PAT 20-APR-2002
ACCESSION		AR184379		
VERSION		AR184379.1	GI:20230344	
KEYWORDS				
SOURCE		Unknown.		
ORGANISM		Unclassified.		
REFERENCE		1 (bases 1 to 2343)		
AUTHORS		Tartaglia, L.A. and Weng, X.		
TITLE		Nucleic acid molecules encoding GLUTX and uses thereof		
JOURNAL		Patent: US 6346374-A 1 12-FEB-2002;		
FEATURES		Location/Qualifiers		
source		1..2343	/organism="unknown"	
			/mol_type="unassigned DNA"	
ORIGIN				
Query Match		100.0%;	Score 2343; DB 6; Length 2343;	
Best Local Similarity		100.0%;	Pred. No. 0;	
Matches 2343;		Conservative 0;	Mismatches 0; Indels 0; Gaps 0;	
QY		1	TCGACCCACGCGTCCGGCCTTGGCAGAGCTGTGGGTCCCTGGACTGAGCCATCAGCTGGG	60
Db		1	TCGACCCACGCGTCCGGCCTTGGCAGAGCTGTGGGTCCCTGGACTGAGCCATCAGCTGGG	60
QY		61	TCACTGAGACCCATGCGCAAGGAAACAATAAGGAATCCAAAGGAATCGGGCCTAGTTCOC	120
Db		61	TCACTGAGACCCATGCGCAAGGAAACAATAAGGAATCCAAAGGAATCGGGCCTAGTTCOC	120
QY		121	CTCAGATGACACGAGCCACGCGGGCCTTCAGGGCCAGGAGGACATGCTGAGATGT	180
Db		121	CTCAGATGACACGAGCCACGCGGGCCTTCAGGGCCAGGAGGACATGCTGAGATGT	180
QY		181	GACCACCTGAGGAGTCGGGTGCGAGGTGCGAAGGAGGAGGACTGGTCTGCTCGCTC	240
Db		181	GACCACCTGAGGAGTCGGGTGCGAGGTGCGAAGGAGGAGGACTGGTCTGCTCGCTC	240
QY		241	CTCGTGGCCCTCCCTCGCGGGCGCCCTTCGGCTCTCCCTCTTCTACGGCTACACCTGTGG	300
Db		241	CTCGTGGCCCTCCCTCGCGGGCGCCCTTCGGCTCTCCCTCTTCTACGGCTACACCTGTGG	300
QY		301	GTGGTGAATGCCCCACCCCGTACATCAAGGCCCTTTACAATGAGTCATGGGAAAGAAGG	360
Db		301	GTGGTGAATGCCCCACCCCGTACATCAAGGCCCTTTACAATGAGTCATGGGAAAGAAGG	360
QY		361	CATGGAGCTCCAATAGACCCAGACACTCTGACTCTGTCTGTGGTCTGTGACTGTGTCCATA	420

361	Db		CATGGAAGTCCAAATAGACCCAGACACTCTGA	CTCTGCTCTGCTGTGACTGTGTCCATA	420
421	Qy		TTCCGCCATCCGTTGAGCTTTGGGGAGCTTAA	TTGTGTGAAGATGATTTGGAAAGGTTCTTTGGG	480
421	Db		TTCCGCCATCCGTTGAGCTTTGGGGAGCTTAA	TTGTGTGAAGATGATTTGGAAAGGTTCTTTGGG	480
481	Qy		AGGAAGCACACTTTGTCTGGCCAAATATAGGGT	TTGCAATTTCTGTGTCATGCTGATGCC	540
481	Db		AGGAAGCACACTTTGTCTGGCCAAATATAGGGT	TTGCAATTTCTGTGTCATGCTGATGCC	540
541	Qy		TGCTCGCTCCAGCCAGGAGCCTTGAAATGCTCAT	TGTGGAGACGCTTCATCATGGGCATA	600
541	Db		TGCTCGCTCCAGCCAGGAGCCTTGAAATGCTCAT	TGTGGAGACGCTTCATCATGGGCATA	600
601	Qy		GATGAGGCGTCCGCTCTCAGTGTGCTCCCATGT	ACCTCAGTGAGATCTCACCCCAAGGAG	660
601	Db		GATGAGGCGTCCGCTCTCAGTGTGCTCCCATGT	ACCTCAGTGAGATCTCACCCCAAGGAG	660
661	Qy		ATCCGTTGCTCTCTGGGGCAGGTGACTGCCAT	CTTTATCTGCATTTGGGTGTTCACTGGG	720
661	Db		ATCCGTTGCTCTCTGGGGCAGGTGACTGCCAT	CTTTATCTGCATTTGGGTGTTCACTGGG	720
721	Qy		CAGCTTCTGGGCCCTGCCAGCTGTGCGAAAG	GAGAGTACTTGGCCCATACCTGTTTGG	780
721	Db		CAGCTTCTGGGCCCTGCCAGCTGTGCGAAAG	GAGAGTACTTGGCCCATACCTGTTTGG	780
781	Qy		GTGATTTGTTGCCCTCGCTTTCTCAGCTGTGT	AGCTTTCCTTTCTCCCGGACAGCCCA	840
781	Db		GTGATTTGTTGCCCTCGCTTTCTCAGCTGTGT	AGCTTTCCTTTCTCCCGGACAGCCCA	840
841	Qy		CGCTACCTGCTCTTTGGAGAAGCACACGAG	CGACAGCTGTGAAGCCTTCCAAAGTTTC	900
841	Db		CGCTACCTGCTCTTTGGAGAAGCACACGAG	CGACAGCTGTGAAGCCTTCCAAAGTTTC	900
901	Qy		TTGGGTTAAACAGACAGCTTTTCCAAAGAG	TAGAGAGTCTTGGCTGAGAGCCACGTGCAG	960
901	Db		TTGGGTTAAACAGACAGCTTTTCCAAAGAG	TAGAGAGTCTTGGCTGAGAGCCACGTGCAG	960
961	Qy		AGGAGCATCCGCTGCTGTCTCGTGTGGAGT	GTCTGAGAGCTTCCCTACGTCCTGCGAG	1020
961	Db		AGGAGCATCCGCTGCTGTCTCGTGTGGAGT	GTCTGAGAGCTTCCCTACGTCCTGCGAG	1020
1021	Qy		GTGCTCACCGTGATTTCTACCACTGGCCCT	GTCTACCACTGTGTGGCCCTCAATTTGG	1080
1021	Db		GTGCTCACCGTGATTTCTACCACTGGCCCT	GTCTACCACTGTGTGGCCCTCAATTTGG	1080
1081	Qy		TTCTATACCAACAGCATCTTTGGAAAGAG	TGGATTCCTCGGCAAGATCCCATACGTC	1140
1081	Db		TTCTATACCAACAGCATCTTTGGAAAGAG	TGGATTCCTCGGCAAGATCCCATACGTC	1140
1141	Qy		ACCTTGAGTACAGGGGACATCGAGACTTTGG	CTGCGCTTCTCTGTTTGGTCAATTGAG	1200
1141	Db		ACCTTGAGTACAGGGGACATCGAGACTTTGG	CTGCGCTTCTCTGTTTGGTCAATTGAG	1200
1201	Qy		CACCTGGGACGGAGACCCCTCTCATTTGGT	GGCTTTTGGGCTCATTGGGCCCTTCTTTGGG	1260
1201	Db		CACCTGGGACGGAGACCCCTCTCATTTGGT	GGCTTTTGGGCTCATTGGGCCCTTCTTTGGG	1260
1261	Qy		ACCTTCACCATCAGCGTGAACCTGTGAGACA	CCAGCCCTCGGGTCCCTTACCTGAGTATC	1320
1261	Db		ACCTTCACCATCAGCGTGAACCTGTGAGACA	CCAGCCCTCGGGTCCCTTACCTGAGTATC	1320
1321	Qy		GTGGGCATTTTGGCCATCATCGCTCTTTTCT	GTCAAGTGGGCGAGGTGGCATCCCGTTCA	1380
1321	Db		GTGGGCATTTTGGCCATCATCGCTCTTTTCT	GTCAAGTGGGCGAGGTGGCATCCCGTTCA	1380
1381	Qy		TTGACCTGGTGAAGTTCTTCAGCAATCTCAG	GGGCGGCTGCCCTTCACTATGAGGACCC	1440
1381	Db		TTGACCTGGTGAAGTTCTTCAGCAATCTCAG	GGGCGGCTGCCCTTCACTATGAGGACCC	1440
1441	Qy		GTCAACTGGCTCTCCAACTTTTCTGTTGGG	CTCCTCTTCCCATTTCAATTAGAAAATCTG	1500



source	1. .2343	Query Match	100.0%	Score 2343	DB 6	Length 2343	0	Gaps	0
		Best Local Similarity	100.0%	Pred. No. 0					
		Matches 2343	Conservative	0	Mismatches	0	Indels	0	
ORIGIN									
Qy	1	TCGACCCACGGTCCGGCTTGGCAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG	60						
Db	1	TCGACCCACGGTCCGGCTTGGCAGTCTGGGGTCCCTGGACTGAGCCATCAGCTGGG	60						
Qy	61	TCACGTAGACCCATGGCAAGGAACAAATAGGAATTCGAAGGAACCTGGGCTAGTTCCC	120						
Db	61	TCACGTAGACCCATGGCAAGGAACAAATAGGAATTCGAAGGAACCTGGGCTAGTTCCC	120						
Qy	121	CTCAGATGACACGAGCCACGGGCTTCCAGGCGCAGGAGGCACTGCTGGAGTGT	180						
Db	121	CTCAGATGACACGAGCCACGGGCTTCCAGGCGCAGGAGGCACTGCTGGAGTGT	180						
Qy	181	GACACCTGAGAGTGGGGTCCAGGTGGAAGGAGAGAGGACTGGTCTGCTCGCTC	240						
Db	181	GACACCTGAGAGTGGGGTCCAGGTGGAAGGAGAGAGGACTGGTCTGCTCGCTC	240						
Qy	241	CTGCTGGGCTTCCCTCGGGGCGCTTGGGCTCCTCTCTACGGCTACCAACCTGTCG	300						
Db	241	CTGCTGGGCTTCCCTCGGGGCGCTTGGGCTCCTCTCTACGGCTACCAACCTGTCG	300						
Qy	301	GTGCTGAATGCCGCCACCCGCTACATCAAGGCCCTTTTACAATGAGTCAATGGAAGG	360						
Db	301	GTGCTGAATGCCGCCACCCGCTACATCAAGGCCCTTTTACAATGAGTCAATGGAAGG	360						
Qy	361	CATGACCTGCAATAGACCCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTCCATA	420						
Db	361	CATGACCTGCAATAGACCCAGACACTCTGACTCTGCTCTGCTGCTGCTGCTCCATA	420						
Qy	421	TTCCGCATCGGTGGACTTGGGGACGTTAATGGAAGATGATGGAAGGTTCTTGGG	480						
Db	421	TTCCGCATCGGTGGACTTGGGGACGTTAATGGAAGATGATGGAAGGTTCTTGGG	480						
Qy	481	AGGAAGCACATTTCTCGGCCAATTAATGGGTTTGCAATTTCTGCTGCTGCTGCTGCTG	540						
Db	481	AGGAAGCACATTTCTCGGCCAATTAATGGGTTTGCAATTTCTGCTGCTGCTGCTGCTG	540						
Qy	541	TGCTCGCTCCAGGACGAGCCCTTTGAAATGCTCAATGTTGGACGCTTCAATCATGGGATA	600						
Db	541	TGCTCGCTCCAGGACGAGCCCTTTGAAATGCTCAATGTTGGACGCTTCAATCATGGGATA	600						
Qy	601	GATGAGGCGTCCGCTCAGTGTGCTCCCATGTAACCTCAGTGAATCTCACCAAGGAG	660						
Db	601	GATGAGGCGTCCGCTCAGTGTGCTCCCATGTAACCTCAGTGAATCTCACCAAGGAG	660						
Qy	661	ATCCGTGGCTCTCTGGGCGAGTGACTGCCATCTTTATCTGCTGCTGGGCTGCTCACTGGG	720						
Db	661	ATCCGTGGCTCTCTGGGCGAGTGACTGCCATCTTTATCTGCTGCTGGGCTGCTCACTGGG	720						
Qy	721	CAGCTTCTGGGCTGCTCCGAGCTGCTGGAAAGAGAGAGTACCTGGCCATACCTGTTTGA	780						
Db	721	CAGCTTCTGGGCTGCTCCGAGCTGCTGGAAAGAGAGAGTACCTGGCCATACCTGTTTGA	780						
Qy	781	GTGATGTGGTCCCTGCGGTTCTCAGCTGTGAGCCCTTCCCTTCTCCCGGACAGCCCA	840						
Db	781	GTGATGTGGTCCCTGCGGTTCTCAGCTGTGAGCCCTTCCCTTCTCCCGGACAGCCCA	840						
Qy	841	CGCTACCTGCTTGGAGAGACACACGAGGCAAGAGCTGTGAAAGCCCTTCCAAACGTTTC	900						
Db	841	CGCTACCTGCTTGGAGAGACACACGAGGCAAGAGCTGTGAAAGCCCTTCCAAACGTTTC	900						
Qy	901	TTGGGTAAAGCAGAGCTTCCCAAGAGGTAGAGGAGGCTCCTGGCTGAGAGCCACGTCGAG	960						
Db	901	TTGGGTAAAGCAGAGCTTCCCAAGAGGTAGAGGAGGCTCCTGGCTGAGAGCCACGTCGAG	960						

Qy	961	AGGAGCATCCGCTCGTGTCCGCTGGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAG	1020
Db	961	AGGAGCATCCGCTCGTGTCCGCTGGAGCTGCTGAGAGCTCCCTACGTCCGCTGGCAG	1020
Qy	1021	GTGGTCAACCGTGATTTGTACCAATGGCCTCTACCACTCTGTGGCTCAATCAATTTGG	1080
Db	1021	GTGGTCAACCGTGATTTGTACCAATGGCCTCTACCACTCTGTGGCTCAATCAATTTGG	1080
Qy	1081	TTCTATACCAACAGCATCTTTGGAAAAGCTGGGATCCCTCCGGCAAAAGATCCCATACGTC	1140
Db	1081	TTCTATACCAACAGCATCTTTGGAAAAGCTGGGATCCCTCCGGCAAAAGATCCCATACGTC	1140
Qy	1141	ACCTTGAGTACAGGGGGCATCGAGATTTGGCTGGCTCTCTCTGTTTGGTCAATTGAG	1200
Db	1141	ACCTTGAGTACAGGGGGCATCGAGATTTGGCTGGCTCTCTCTGTTTGGTCAATTGAG	1200
Qy	1201	CACCTGGGACGAGACCCCTCTCTATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG	1260
Db	1201	CACCTGGGACGAGACCCCTCTCTATTTGGTGGCTTTGGGCTCATGGGCTCTCTTTTGGG	1260
Qy	1261	ACCTCACCATCAGCGTGAACCCCTGCAGGACCAACGCCCCCTGGGTCCCTACCTGAGTATC	1320
Db	1261	ACCTCACCATCAGCGTGAACCCCTGCAGGACCAACGCCCCCTGGGTCCCTACCTGAGTATC	1320
Qy	1321	GTGGGCAITCTGGCCATCATCGCCTCTTCTGAGTGGGCCAGGTGGGATCCGGTTCAAC	1380
Db	1321	GTGGGCAITCTGGCCATCATCGCCTCTTCTGAGTGGGCCAGGTGGGATCCGGTTCAAC	1380
Qy	1381	TTGACTGGTGAATTTCTCCAGCAATCTCAGCGCGGCTGCTTCAATTCAGGACAC	1440
Db	1381	TTGACTGGTGAATTTCTCCAGCAATCTCAGCGCGGCTGCTTCAATTCAGGACAC	1440
Qy	1441	GTCAACTGGCTCTCCAACTTTGCTGGGCTCTCTCTTCCCAATTCATTCAGAAAAGTCTG	1500
Db	1441	GTCAACTGGCTCTCCAACTTTGCTGGGCTCTCTCTTCCCAATTCATTCAGAAAAGTCTG	1500
Qy	1501	GACACCTACTGTTTCTCTAGTCTTTGCTACAAATTTGTATCACAGGTGCTATCTACCTGTAT	1560
Db	1501	GACACCTACTGTTTCTCTAGTCTTTGCTACAAATTTGTATCACAGGTGCTATCTACCTGTAT	1560
Qy	1561	TTTGTCTGCTGAGACCAAAAACAGAACCTATGCGAATCAGCCAGGCAATTTTCCAAA	1620
Db	1561	TTTGTCTGCTGAGACCAAAAACAGAACCTATGCGAATCAGCCAGGCAATTTTCCAAA	1620
Qy	1621	AGGAAACAAAGCATACCCACGAGAGAAAATCGACTCAGCTGTCTACTGATGCTCTGCT	1680
Db	1621	AGGAAACAAAGCATACCCACGAGAGAAAATCGACTCAGCTGTCTACTGATGCTCTGCT	1680
Qy	1681	TCCTCTCTTCTACTACTCCGAATACAGCTGGATTCAGAGTCCGCGCCACCCACCGCC	1740
Db	1681	TCCTCTCTTCTACTACTCCGAATACAGCTGGATTCAGAGTCCGCGCCACCCACCGCC	1740
Qy	1741	ACCAAAAAAGAACACCCCATTTGTAACGGTCAATGGTATTTCTCAACCTGGAATGACCT	1800
Db	1741	ACCAAAAAAGAACACCCCATTTGTAACGGTCAATGGTATTTCTCAACCTGGAATGACCT	1800
Qy	1801	TCCTCTATCTTCTCTCTGGAGAACCAAGTCATGATGTCAGCAAGAGCTTGGATTT	1860
Db	1801	TCCTCTATCTTCTCTCTGGAGAACCAAGTCATGATGTCAGCAAGAGCTTGGATTT	1860
Qy	1861	TGGAGCATGGGTTTGAATTTCCAGTCATTCATTTTATTTTCCAGCAAAATATTTAACAGT	1920
Db	1861	TGGAGCATGGGTTTGAATTTCCAGTCATTCATTTTATTTTCCAGCAAAATATTTAACAGT	1920
Qy	1921	ACTGACATGTCCTATGTTTGTATTTTACCACTGGTTTATCAATGGGAGGAGAGAGAG	1980
Db	1921	ACTGACATGTCCTATGTTTGTATTTTACCACTGGTTTATCAATGGGAGGAGAGAGAG	1980
Qy	1981	AGAGAGAGAGAGAGATGCTATTTAAAGCTTGAAGTCTAGGCTGTGACGGTGGCTC	2040
Db	1981	AGAGAGAGAGAGAGATGCTATTTAAAGCTTGAAGTCTAGGCTGTGACGGTGGCTC	2040
Qy	2041	ACGCTGTATTCACGACACTTTTGGGAGCGGAGGTGGTGGATTCGTGAGGTTCAGGAGATT	2100

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: April 7, 2004, 21:03:37 ; Search time 9011 Seconds  
(without alignments)  
11269.862 Million cell updates/sec  
Title: US-09-981-947B-1  
Perfect score: 2343  
Sequence: 1 tggaccacagctccggcct.....aaaaaaaaagggcgccgcg 2343

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0  
Searched: 3470272 seqs, 21671516995 residues 6940544  
Total number of hits satisfying chosen parameters:  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

GenEmbl:.\*  
1: gb.ba:.\*  
2: gb.htg:.\*  
3: gb.in:.\*  
4: gb.om:.\*  
5: gb.ov:.\*  
6: gb.pat:.\*  
7: gb.ph:.\*  
8: gb.pl:.\*  
9: gb.pr:.\*  
10: gb.ro:.\*  
11: gb.sts:.\*  
12: gb.sy:.\*  
13: gb.un:.\*  
14: gb.vi:.\*  
15: em.ba:.\*  
16: em.fun:.\*  
17: em.hum:.\*  
18: em.in:.\*  
19: em.mu:.\*  
20: em.om:.\*  
21: em.or:.\*  
22: em.ov:.\*  
23: em.pat:.\*  
24: em.ph:.\*  
25: em.pl:.\*  
26: em.ro:.\*  
27: em.sts:.\*  
28: em.un:.\*  
29: em.vi:.\*  
30: em.htg.hum:.\*  
31: em.htg.inv:.\*  
32: em.htg.other:.\*  
33: em.htg.mus:.\*  
34: em.htg.pln:.\*  
35: em.htg.rod:.\*  
36: em.htg.man:.\*  
37: em.htg.vrt:.\*  
38: em.sy:.\*  
39: em.htgo.hum:.\*  
40: em.htgo.mus:.\*  
41: em.htgo.other:.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2343	100.0	2343	6	AR135669 Sequence
2	2343	100.0	2343	6	AR184379 Sequence
3	1651	70.5	2032	6	AX709981 Sequence
4	1646.4	70.3	1863	9	AF210317 Homo sapi
5	1445.2	61.7	1885	9	BC018897 Homo sapi
6	1317.8	56.2	2842	6	AX339083 Sequence
7	1141	48.7	2216	6	AX179745 Sequence
8	1089.2	46.5	2022	10	AF469480 Mus muscu
9	636.6	27.2	160402	2	AC067775 Homo sapi
10	636.6	27.2	165221	9	AC098976 Homo sapi
11	616.2	26.3	3682	9	HSU21164 Human dopam
12	592.2	25.3	597	9	AF421859 Homo sapi
13	523.6	22.3	1296	10	AF490463 Mus muscu
14	515.8	22.0	3243	10	BC006076 Mus muscu
15	427.6	18.3	1539	6	AX753091 Sequence
16	427	18.2	153113	9	AC006453 Homo sapi
17	424.8	18.1	1714	10	RATG75 D28562 Rattus norv
18	424.4	18.1	1539	6	AX686979 Sequence
19	424.4	18.1	2080	10	BC023500 Mus muscu
20	423.4	18.1	2137	10	RATGLUTV L05195 Rattus norv
21	423.4	18.1	138007	2	AC141282 Homo sapi
22	423.4	18.1	173505	2	AC141297 Homo sapi
23	423.4	18.1	213386	2	AC139261 Homo sapi
24	422	18.0	105643	9	AC005629 Homo sapi
25	421.6	18.0	2169	10	RATGLUTS D13871 Rattus norv
26	421.2	18.0	2069	10	AF161071 Mus muscu
27	421	18.0	168754	2	AC027487 Homo sapi
28	419.4	17.9	158052	9	AC027612 Homo sapi
29	419.4	17.9	159575	2	AC141407 Homo sapi
30	419.4	17.9	165225	9	AC116050 Homo sapi
31	419.4	17.9	195985	2	AC141251 Homo sapi
32	419.4	17.9	197109	2	AC141252 Homo sapi
33	416.4	17.8	2213	9	BC001692 Homo sapi
34	416.4	17.8	2213	9	BC001820 Homo sapi
35	416.4	17.8	2218	6	AX332575 Sequence
36	416.4	17.8	2218	6	AX333214 Sequence
37	416.4	17.8	2218	9	HUMGLUT5 M55531 Human glucoc
38	416	17.8	155956	2	AC144898 Homo sapi
39	416	17.8	193913	2	AC145614 Homo sapi
40	414	17.7	57598	9	AL513526 Human DNA
41	410	17.5	7434	6	AX348519 Sequence
42	407.8	17.4	7434	6	AX348520 Sequence
43	402	17.2	2166	4	ECA555215 Equus cab
44	390.8	16.7	2147	4	OAR315928 Ovis arie
45	376	16.0	1607	6	AX556510 Sequence

ALIGNMENTS

RESULT 1  
AR135669  
LOCUS AR135669 2343 bp DNA linear PAT 16-JUN-2001  
DEFINITION Sequence 1 from patent US 6136547.  
ACCESSION AR135669  
VERSION AR135669.1 GI:14476341  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2343)  
AUTHORS Tartaglia,L.A. and Weng,X.  
TITLE Nucleic acid molecules encoding glutx and uses thereof  
JOURNAL Patent: US 6136547-A 1 24-OCT-2000;  
FEATURES Location/Qualifiers



```
Db 394 TACATGAGTCATGGGAAAGAGCATGGACGTCCATAGACCCAGACACTCTGACTCTG 453
Qy 397 CTCTGGTCTGTGACTGTGTCCATATTCCGCATCGGTGGACTTGTGGGACGTTAATTGTG 456
Db 454 CTCTGGTCTGTGACTGTGTCCATATTCCGCATCGGTGGACTTGTGGGACATTAATTGTG 513
Qy 457 AAGATGATTGGAAAGGTTCTTGGGAGGAAAGCACACTTGTCTGGCCCAATAATGGGTTTGCA 516
Db 514 AAGATGATTGGAAAGGTTCTTGGGAGGAAAGCACACTTGTCTGGCCCAATAATGGGTTTGCA 573
Qy 517 ATTTCTGCTGCAATTGTGATGGCTCTCGCTCCAGGACAGGAGCCTTTGAATGCTCATT 576
Db 574 ATTTCTGCTGCAATTGTGATGGCTCTCGCTCCAGGACAGGAGCCTTTGAATGCTCATT 632
Qy 577 GTGGGACGTTTCATCATGGGCAATAGATGGAGGCGTCCCTCAGTGTGCTCCCATGTAC 636
Db 633 GTGGGACGTTTCATCATGGGCAATAGATGGAGGCGTCCCTCAGTGTGCTCCCATGTAC 691
Qy 637 CTCAGTGAGATCTCACCCAGGAGATCCGTGTGCTCTTGGGGCAGGTGACTGCCATCTTT 696
Db 692 CTCAGTGAGATCTCACCCAGGAGATCCGTGTGCTCTTGGGGCAGGTGACTGCCATCTTT 750
Qy 697 ATCTGCAATTGGCG 709
Db 751 ATCTGCAATTGGGG 763
```

Search completed: April 8, 2004, 16:08:51  
Job time : 5995 secs